NEW YORK UNIVERSITY
INSTITUTE OF
MATHEMATICAL SCIENCES

Integration of Functionals

K. O. FRIEDRICHS
H. N. SHAPIRO
ET AL

WILBUR L. CROSS LIBRARY
UNIVERSITY OF CONNECTICUT

QA
320
F75

Digitized by the Internet Archive
in 2012 with funding from
LYRASIS members and Sloan Foundation

http://archive.org/details/seminarintegrati00frie

SEMINAR

Integration of Functionals

K. O. Friedrichs

H. N. Shaprio et al

New York University

Institute of Mathematical Sciences

1957

QA
320
F75

Table of Contents

Chapter	Title
I	Integration of functionals. (K. O. Friedrichs)
II	Remarks on the quantum theory of fields. (K. O. Friedrichs)
III	Compatible measures. (H. N. Shapiro)
IV	Integration spaces. (H. N. Shapiro)
V	Integration over Hilbert space. (H. N. Shapiro)
VI	A property of Gaussian measures. (H. N. Shapiro)
VII	The integration of polynomial functionals. (K. O. Friedrichs)
VIII	Integration over Hilbert space and related norms. (J. Schwartz)
IX	The Wiener integral. (H. N. Shapiro)
X	Remarks about the Wiener measure. (K. O. Friedrichs)
XI	Stochastic processes. (B. Wendroff)
XII	Integration over outer Hilbert space. (H. N. Shapiro
XIII	The Integrability of Polynomials (T. Seidman)
XIV	The Feynman-Kac formula. (J. Schwartz)
XV	The Wiener-Siegel causal theory of quantum mechanics. (J. Schwartz)
B1-10	Bibliography - Appendix
1 - V111	Errata
E.1-3	Epilogue: Remarks on Theorems by Paley and Wiener

PREFACE

The present notes contain a report of lectures given in a seminar on Functional Analysis, held in the Spring of 1955 at the Institute of Mathematical Sciences.

The general topic of this seminar was "integration in functional space", with the main emphasis placed upon the integration of functionals defined over a Hilbert space. The first chapters (I, II, VI, VIII) of the lectures consist of an elaboration of an approach to this problem indicated in an earlier work of one of the contributors. In pursuing this, a result on the invariance of integrals of cylinder functionals was obtained, independently and at the same time, by J. Schwartz and by K.O. Friedrichs and H.N. Shapiro (Chapters VI, VII). This result in turn led to an extension of the Gaussian measure over Hilbert space into an "outer Hilbert Space" in which the extended measure is totally additive. (Chapter XII*)

The notes include some discussion (Chapters IX, X) of the structural relationship of the integral considered here to the Wiener integral. It should also be noted that Bochner's extension of the Kolmogoroff Theorem** is similar in spirit to the ideas developed here. However, Bochner's Theorem does not cover the situation considered in the notes.

In the present notes no discussion is included of the work of I.E. Segal in this area.

The seminar was directed by K.O. Friedrichs and H.N. Shapiro and assisted by Thomas Seidman and Charles Lytle***.

---

* A preliminary report on these results has appeared in the Proceedings of the National Academy of Sciences, Vol. 43, No. 4, pp. 336-338, April 1957.

** S. Bochner, Harmonic Analysis and the Theory of Probability, Chapter V. University of California Press, 1955.

*** This work was supported by the Office of Naval Research and the National Science Foundation.

Chapter I. Integration of Functionals

The question I should like to pose and which we shall consider in the present seminar is this: how does one integrate functionals and what motivates the attempt?

Several approaches to the integration of functionals have been developed in the course of many years by Gateaux, Levy, Daniell, Wiener, Jessen, Kolmogoroff and others. We intend to study some of these approaches but each of them seems a little too restricted in some respects. These approaches yield integrals which are most suitable for the purposes of probability theory but, without modification, they do not seem to be suitable for use in quantum theory.

An explanation of the need for the integration of functionals in the quantum theory of fields will be given in the second chapter. First, however, I should like to present one possible way of defining an integral of functionals which would seem to be wide enough to include both those integrals needed in probability and those needed in quantum theory.

The functionals we shall integrate will be functions defined over a class of functions, specifically functions f of a variable ξ where the values taken on by ξ are functions of a variable s. We take s as a variable over a given base space $\mathcal{S}$, with a given measure differential, $\mu(ds)$. For concreteness one may think of s as a single real variable running over the s-axis and in our description we shall take $\mathcal{S}$ as the real axis. The class $\{\xi\}$ of independent functions ξ will, at present, be left unspecified except that we require that it include all piecewise constant functions of s. The functions in $\{\xi\}$ will be denoted by '$\xi(\cdot)$' or more simply by 'ξ', while by '$\xi(s)$' we denote the value this function assumes when the independent variable takes on the particular value s. The functional will be denoted by '$f[\cdot]$' or by 'f' and its value for the particular function, $\xi(\cdot)$, will be denoted by $f[\xi(\cdot)]$ or $f[\xi]$. The class of these functionals will also be left temporarily unspecified.

We shall first define the integral for a functional f of a very special type.

Let us "partition" $\mathcal{S}$, expressing it as the sum of a finite number of regions. This partition of $\mathcal{S}$ will be denoted by '$\mathcal{P}$'. Considering s as a single real variable, such a partition may be given by subdividing the s-axis $\mathcal{S}$ into intervals, $\mathcal{I}_\nu = (s_{\nu-1}, s_\nu)$ $(\nu = 1,\ldots n)$, with $s_o = -\infty$ and $s_n = +\infty$. We consider functions $\xi(\cdot)$ which are constant in each of the subregions $\mathcal{I}_\nu$; that is

$$\hat{\xi}(s) = c_\nu \qquad \text{for } s \text{ in } \mathcal{I}_\nu \quad (\nu = 1,\ldots n).$$

Other functions ξ may be projected into such piecewise constant functions. To this end one employs the measure μ and sets

$$\xi_\nu = \frac{1}{\Delta_\nu} \int_\nu \xi(s)\mu(ds) \quad \text{with } \Delta_\nu = \mu(\mathcal{I}_\nu).$$

The projection of the function ξ thus described will be denoted by $P\xi$. We have

$$P\xi(s) = \xi_\nu \qquad \text{for } s \text{ in } \mathcal{I}_\nu \quad (\nu=1,\ldots n).$$

Evidently, the projection P depends on the partition $\mathcal{P}$ of $\mathcal{S}$ into the subregions $\mathcal{I}_\nu$.

We now consider a special class, $\{\hat{f}\}$, of functionals $\hat{f}$ which depend on their argument ξ only in as much as they depend on $P\xi$. More precisely, these are the functionals such that, for each $\xi(\cdot)$ in the (unspecified) class of functions,

$$\hat{f}[\xi] = \hat{f}[P\xi]$$

Such a functional will be called "compatible with the projection P". In view of the custom in the case of a discrete measure μ, a functional compatible with some projection will also be called a "cylinder functional". As a nontrivial example of such a special functional, note that

$$f[\xi] = \int_{\mathcal{S}} \xi(s)ds = \sum_\nu \int_{\mathcal{S}_\nu} \xi(s)ds$$

has this property of compatibility for any subdivision of $\mathcal{S}$.

Evidently, a cylinder functional can be completely described with the aid of a function of n variables, namely by a function $f_P(x_1,\ldots x_n)$ such that,

$$f[\phi] = f[P\phi] = f_P(\phi_1,\ldots\phi_n)$$

where $$[P\phi](s) = \phi_\nu \quad \text{for sin} \mathcal{S}_\nu \quad (\nu=1,\ldots n).$$

The argument, $(\phi_1,\ldots\phi_n)$, of $f_P(\cdot)$ is a set of n numerical variables corresponding to the argument, $\phi = \phi(\cdot)$, of $f[\cdot]$ in a way determined by the projection P.

For the function f_P, as a function of a finite number of variables, there is an integration theory and we should like to define the integral of f in terms of an integral of f_P. Thus, we integrate f_P over the whole $(\phi_1,\ldots,\phi_n)$ - space with respect to measures m_ν which must be properly chosen subject to conditions which we shall discuss. Accordingly, we proceed to consider the integral

$$\int\cdots\int f_P(\phi_1\ldots\phi_n)\, dm_1(\phi)\ldots dm_n(\phi_n),$$

the n-tuple integral being extended over the whole $(\phi_1\ldots\phi_n)$-space. This integral will be considered as the integral of the functional $f[\cdot]$ over the space of functions $\{\phi\}$.

At present we denote this integral $I_P(f)$. I.e. - we define

$$I_P(f) = \int^n f_P(\phi_1 \ldots \phi_n) \prod_{\nu=1}^{n} d\mu_\nu(\phi_\nu).$$

As defined, the integral $I_P(f)$ depends on the projection P. We ask: under which circumstances is $I_P(f)$ independent of the choice of P?

Let $\mathcal{P}'$ be a refinement of the partition $\mathcal{P}$ and let P' be the corresponding projection. If, first, the fine projection P' is applied to a function ϕ and then the coarser projection P is applied to the result the final result is the same as if P had been applied directly to ϕ. This is so because of the special nature of the projections P and P' which we have adopted. We express the fact by the formula,

$$PP'\phi = P\phi \quad \text{or} \quad PP' = P.$$

We call a projection P' such that PP' = P, a <u>sub-projection</u> of P. We have just seen that a refined partition generates a sub-projection.

We note that if f is a functional compatible with a projection P it is compatible with each subprojection P' as

$$f[\phi] = f[P\phi] = f[PP'\phi] = f[P'\phi],$$

the last equality being the same as the first with ϕ replaced by $P'\phi$, the second following from the formula, PP' = P.

As a consequence of this fact the integral $I_{P'}(f)$ is defined. Naturally, we seek conditions assuring that the integral $I_P(f)$ is not changed if $\mathcal{P}$ is replaced by a refinement $\mathcal{P}'$:

$$I_P(f) = I_{P'}(f) \quad .$$

As $f[P'\phi] = f[P\phi] = f[\phi]$, these conditions will be conditions on the measure m_ν. In order to formulate such conditions let us assume that for each ν the measure m_ν depends only on $\mathcal{J}_\nu$ and

not on the whole partition $\mathcal{P}$. It is then sufficient to consider a single segment $\mathcal{I}_0$ of $\mathcal{P}$ which is divided into two segments $\mathcal{I}_1$ and $\mathcal{I}_2$ by $\mathcal{P}$. Let dm_0, dm_1 and dm_2 be the corresponding measure differentials to be used in defining $I_p(f)$ and $I_{p'}(f)$. Then, an easy argument shows, the requirement can be expressed by the condition that the identity,

$$\iint \chi\left(\frac{\phi_1 \Delta_1 + \phi_2 \Delta_2}{\Delta_1 + \Delta_2}\right) \, dm_1(\phi_1) dm_2(\phi_2) = \int \chi(\phi_0) dm_0(\phi_0),$$

holds for arbitrary functions $\chi(\cdot)$ where, as before,

$$\Delta_\nu = \mu(\mathcal{I}_\nu) = \int_{\mathcal{I}_\nu} \mu(ds).$$

If the measures m_ν satisfy this condition they will be called <u>compatible</u>. Examples of such compatible measures will be given a little later. We shall refer to this identity as the <u>compatibility condition</u>.

We shall assume from now on that the measures m_ν $(\nu=1,\ldots n)$ are compatible. In that case the integral $I_{p'}(f)$ of any functional f which is compatible with a projection P exists and is independent of the choice of the subprojection P'. We express this by omitting the subscript, -

$$I(f) = I_{p'}(f) = I_p(f) = \int^n f_p(\phi,\ldots\phi_n) \prod_{\nu=1}^{n} dm_\nu(\phi_\nu).$$

It should also be noted that if f is compatible with two projections, P and $\hat{P}$, arising from different partitions, $\mathcal{P}$ and $\hat{\mathcal{P}}$, then there is a refinement $\mathcal{P}'$ common to $\mathcal{P}$ and $\hat{\mathcal{P}}$ and generating a projection P', It follows that

$$I(f) = I_p(f) = I_{p'}(f) = I_{\hat{p}}(f).$$

It is important to note that in order to show this we have needed the fact that the measure μ which was associated with $\mathcal{S}$ is held fixed throughout. Consequently, for every cylinder functional f(i.e. one compatible with some projection) the integral I(f) exists and is independent of the projection.

Up to now it has been assumed that f was a cylinder functional. We proceed to extend the class of functionals for which the integral is defined. We note that the cylinder functionals form a linear space normed by

$$\|f\| = I(|f|) = \int^{n} |f_P(\phi_1 \ldots \phi_n)| \, dm_1(\phi) \ldots dm_n(\phi_n).$$

Clearly, if f is compatible with P and α is a number, then αf is compatible with P ; -

$$(\alpha f)[\phi] = \alpha f[\phi] = \alpha f[P\phi] = (\alpha f)[P\phi].$$

Also, if f is compatible with P and g with $\hat{P}$, then (f+g) is compatible with P', the common subprojection of P and $\hat{P}$; -

$$(f+g)[\phi] = f[\phi]+g[\phi] = f[P\phi]+g[P\phi] = f[\hat{P}'\phi]+g[P'\phi] = (f+g)[P'\phi].$$

Thus all finite linear combinations of cylinder functionals are again cylinder functionals. In fact, an arbitrary function of a finite number of cylinder functionals is a cylinder functional and, in particular, if f is a cylinder functional so is $|f|$. As $I(|f|) = I_p(|f|)$ for some P and as $I_p(|\cdot|)$ is a finite-dimensional integral known to be positive-definite, subadditive, and absolutely multiplicative, $I(|\cdot|)$ is seen to be a norm on the space of cylinder functionals. The space will not, in general, be a Banach space and we extend it to a Banach space by the usual method for completion of a metric space by the addition of ideal elements.

That is ; - we consider sequences of cylinder functionals, $\{f_n\}$ (n=1,2,...), such that,

$$I(|f_m - f_n|) \longrightarrow 0 \quad \text{as } m,n \longrightarrow \infty .$$

To each such Cauchy sequence f_n we may — as is well known — assign an "ideal" functional $\hat{f}$, with the stipulation that the same ideal functional will be assigned to two Cauchy sequences if and only if they are equivalent. Two sequences, $\{f_n\}$ and $\{f_n^*\}$, are equivalent if $I(|f_n - f_n^*| \longrightarrow 0$ as $n \longrightarrow \infty$. If $\{f_n\}$ is a Cauchy sequence of cylinder functionals to which there corresponds an ideal functional f then $I(f_n)$ converges to a limit which we call the integral of f ; -

$$I(f) = \lim_{n \longrightarrow \infty} I(f_n).$$

It is clear that I(f) is independent of the particular equivalent soquence used in its definition. If an ideal element $\hat{f}$ corresponds to a class of equivalent Cauchy sequences among which is one $\{f_n\}$ for which $f_n = f$ for a certain cylinder functional f and each n = 1,2,..., then we identify $\hat{f}$ with f. Clearly $I(\hat{f}) = I(f)$.

We have thus defined the integral of ideal functionals, which, while not necessarily themselves cylinder functionals, can be approximated in the mean, i.e.-with respect to the norm $I(|\cdot|)$, by cylinder functionals. Let us consider the situation involved in integrating such a functional. Take a sequence $\{f_n\}$ of approximating cylinder functionals where each f_n is compatible with the projection P_n generated by the partition $\mathcal{P}_n$. Let $\tilde{\mathcal{P}}_k$ be the common refinement of $\mathcal{P}_1 \ldots \mathcal{P}_k$ and let $\tilde{P}_k$ be the corresponding projection. Then, cutting off the sequence $I(f_n)$ at the k-th step, one has

$$(*) \quad I(f_n) = \int^{n_k} (f_n)_{\tilde{P}_k} (\phi_1 \ldots \phi_{n_k}) \prod_{\nu=1}^{n_k} dm_\nu(\phi_\nu) \quad (n=1 \ldots k)$$

where, for each n, the integral is taken with respect to the same projection $\tilde{P}_k$. In general, of course, $n_k \longrightarrow \infty$ as $k \longrightarrow \infty$.

In order to remind one of the fact that the integral I(f) has resulted from an integral of the form (*) we adopt the notation,

$$I(f) = \int^{\infty} f[\phi] \prod_s dm_s(\phi(s)),$$

where s is a continuous index varying over the entire base space $\mathcal{S}$ and, for each s, the quantity $\phi(s)$ is an independent variable which we integrate from $-\infty$ to $+\infty$.

Heuristically one might imagine that "in the limit" there would be a "partition" $\tilde{\mathcal{P}}$ (giving $\mathcal{S}$ as the sum of an infinite number of primitive sets consisting of a single point each) generating a "projection" $\tilde{P}$ with respect to which the limiting integral might be taken. The projection $\tilde{P}$ is, then, to be regarded as one which sends a function ϕ into the set of its values $\phi(s)$.

The true meaning of the integral does not lie in this heuristic description but is, of course, given by the limiting procedure with approximating cylinder functionals described above. The notation, however, is suggestive of various formal manipulations which we shall investigate later on.

Before raising further questions I shall describe a special set of measure differentials, the Gaussian measure differentials, which are compatible in the sense explained above;

$$dm_\nu(\phi_\nu) = e^{-\phi_\nu^2 \Delta_\nu} \sqrt{\frac{\Delta_\nu}{\pi}}\, d\phi_\nu \ ,$$

The compatibility is well-known and easily proved - I shall not present such a proof here. In the following I shall adopt this measure unless otherwise stated. Using this measure we shall employ a different integration notation. In the form shown for the integration of ideal functionals we replace Δ_ν by $\mu(ds)$ and $d\phi_\nu$ by $\delta\phi(s)$. Altogether we write

$$I(f) = \int^{\infty} f[\phi] \prod_s e^{-\phi^2(s)\mu(ds)} \sqrt{\frac{\mu(ds)}{\pi}}\ \delta\phi(s)$$

We shall go a step further, since

$$\prod_\nu e^{-\phi_\nu^2 \Delta_\nu} = e^{-\sum_\nu \phi_\nu^2 \Delta_\nu}$$

we write symbolically

$$I(f) = \int^{\infty} f[\phi] e^{-\int_S \phi^2(s)\mu(ds)} \prod_s \sqrt{\frac{\mu(ds)}{\pi}}\ \delta\phi(s).$$

Employing the expression $\int_S \phi^2(s)\mu(ds)$ symbolically is not, of course, meant to imply that this integral is finite for the functions $\{\phi(\cdot)\}$ considered. I will come back to this point when we shall discuss the relationship of this integral to the Wiener integral.

It is frequently desirable to work with a <u>translation invariant</u> integral; i.e. - with an integral $J(\cdot)$, defined for functionals $F[\cdot]$, for which the relation

$$J(F) = J(TF)$$

holds where T is an operator of the form

$$TF[\phi] = F[\phi + \phi_0]$$

for any arbitrary fixed function ϕ_0. The integral $I(\cdot)$ does not have this property - but it can be modified to a translation invariant integral in an obvious way. To this end we associate functionals f and F such that

$$F[\phi] = f[\phi] e^{-\int_S \phi^2(s)\mu(ds)}$$

and we introduce an operator $\tilde{T}$ such that

$$\tilde{T}f[\phi] = e_0 f[\phi + \phi_0] e^{-2\int_S \phi(s)\phi_0(s)\,\mu(ds)}$$

where $$c_0 = \exp[-\int_{\mathcal{S}} \phi_0^2(s)\mu(ds)].$$

If $\tilde{F}$ corresponds to $\tilde{Tf}$ as F corresponds to f it is clear that

$$\tilde{F}[\phi] = (\tilde{Tf}[\phi])e^{-\int_{\mathcal{S}} \phi^2(s)\mu(ds)}$$

$$= f[\phi + \phi]e^{-\int_{\mathcal{S}} [\phi(s) + \phi_0(s)]^2\mu(ds)}$$

$$= F[\phi + \phi_0] = TF[\phi] \quad .$$

If, for any F, we now define J(F) by J(F) = I(f) where f corresponds as above to F then the integral $J(\cdot)$ is translation invariant, i.e. -

$$J(\tilde{F}) = J(F)$$

as $$I(\tilde{Tf}) = I(f).$$

This last formula is easily shown for cylinder functionals f - note that if f is a cylinder functional so is $\tilde{Tf}$ -- and for non-cylinder functionals we note that, obviously, if f is integrable so is $\tilde{Tf}$ and the equality of the integrals holds in the limit as it holds for each approximating cylinder functional.

We shall now describe functionals f or <u>classes of functionals</u> for which the integral I(f) exists. The simplest such example is given by the quadratic form,

$$f[\phi] = \iint_{\mathcal{S}} b(s',s'')\phi(s')\phi(s'')\mu(ds')\mu(ds''),$$

where the double integral is taken over the Cartesian product of $\mathcal{S}$ with itself. This functional is compatible with a partition, $\mathcal{P}$, if b(s',s") is constant on each "rectangle" $\mathcal{I}_{\nu'} \times \mathcal{I}_{\nu''}$ ($\nu', \nu'' = 1,2...n$). In general such a functional f[•] can be approximated by cylinder functionals of the same form and, therefore, the procedure for evaluating the integral I(f) is well-defined. Instead of computing I(f) by strictly following the definition, we shall evaluate it by a formal procedure suggested by the notation, computing as if "$\mu(ds)$" were a number

and as if the sign "$\int$" stood for a summation. In this way we find

$$I(f) =$$

$$\int_{\mathcal{S}'\mathcal{S}''} \left\{ \iint b(s',s'')\phi(s')\phi(s'')\mu(ds')\mu(ds'') \right\} \prod_s e^{-\phi^2(s)\mu(ds)} \sqrt{\frac{\mu(ds)}{\pi}}\, \delta\phi(s).$$

As we are proceeding on a purely formal basis we may, without further justification, interchange the order of integration bringing the infinitely iterated integral with respect to the variables $\{\phi(s)\}$ inside the integration with respect to s' and s" and at the same time bringing the term b(s',s") outside the integral, $\int^{\infty} \prod_s \delta\phi(s)$, with respect to which it is constant. Thus we get

$$I(f) =$$

$$\iint \left\{ \int^{\infty} b(s',s'')\phi(s')\phi(s'') \prod_s e^{-\phi(s)\mu(ds)} \sqrt{\frac{\mu(ds)}{\pi}}\, \delta\phi(s) \right\} \mu(ds')\mu(ds'')$$

$$= \iint b(s',s'')i(s',s'')\mu(ds')\mu(ds'')$$

where

$$i(s',s'') = \int^{\infty} \phi(s')\phi(s'') \prod_s e^{-\phi^2(s)\mu(ds)} \sqrt{\frac{\mu(ds)}{\pi}}\, \delta\phi(s).$$

Let us note here the three identities;

(1) $$\int_{-\infty}^{\infty} e^{-x^2/\sigma} \frac{1}{\sigma\sqrt{\pi}}\, dx = 1 \quad (\sigma > 0),$$

(2) $$\int_{-\infty}^{\infty} x e^{-x^2/\sigma} \frac{1}{\sigma\sqrt{\pi}}\, dx = 0, \text{ and}$$

(3) $$\int_{-\infty}^{\infty} x^2 e^{-x^2/\sigma^2} \frac{1}{\sigma\sqrt{\pi}}\, dx = \sigma^2 \quad (\sigma > 0).$$

Using them we may evaluate $i(s',s'')$. Let $\mu(ds) = \sigma^{-2}$. Then we have

$$i(s',s'') = \int^{\infty} \phi(s')\phi(s'')\prod_s e^{-\phi^2(s)/\sigma^2} \frac{1}{\sigma\sqrt{\pi}} \delta\phi(s).$$

We consider first the case: $s' \neq s''$. Here

$$i(s',s'') = \prod_{s\neq s',s''}\left[e^{-\phi^2(s)/\sigma^2} \frac{1}{\sigma\sqrt{\pi}} \delta\phi(s)\right]$$

$$\left[\int \phi(s')e^{-\phi^2(s')/\sigma^2} \frac{1}{\sigma\sqrt{\pi}} d\phi(s')\right]\cdot$$

$$\cdot\left[\int \phi(s'')e^{-\phi^2(s'')/\sigma^2} \frac{1}{\sigma\sqrt{\pi}} d\phi(s'')\right]$$

where we have considered " $\int^{\infty} \ldots\prod\ldots\delta\phi(s)$" as an iterated integral and have separated out the integrations with respect to independent variables. We note that, by (1), each of the integrals forming the infinite product equals one and, by (2), the integrals giving the last two factors both vanish. Hence

$$i(s',s'') = 0 \qquad \text{for } s' \neq s''.$$

We now consider the remaining case: $s' = s''$. Here

$$i(s',s') = \prod_{s\neq s'}\left[\int e^{-\phi^2(s)/\sigma^2} \frac{1}{\sigma\sqrt{\pi}} d\phi(s)\right]\cdot$$

$$\cdot\left[\int \phi^2(s')e^{-\phi^2(s')/\sigma^2} \frac{1}{\sigma\sqrt{\pi}} d\phi(s')\right].$$

Again the infinite product equals one but now, by (3), the second factor equals σ^2. Hence

$$i(s',s') = \frac{1}{\mu(ds')}$$

Thus we find

$$I(f) = \iint b(s',s'')i(s',s'')\mu(ds')\mu(ds'')$$

$$= \iint_{s'=s''} b(s',s'') \frac{1}{\mu(ds')} \mu(ds') \mu(ds'')$$

and hence

$$I(f) = \int_S b(s,s)\mu(ds).$$

We may say that the value of the integral of the functional is the trace of the integral operator with the kernel $b(s',s'')$. The occurence of such a trace is very typical for evaluation of the integral of polynomials with respect to Gaussian measures. I shall discuss this in more detail later on; see Chapter VIII.

In a similar way we compute the integral I(f) for the functional

$$f[\phi] = e^{\int_S b(s)\phi(s)\,\mu(ds)}$$

From the identity

$$\int e^{(bx-x^2)/\sigma} \frac{1}{\sigma\sqrt{\pi}}\, dx = e^{b^2/4\sigma^2}$$

we have, again letting $\mu(ds) = \sigma^{-2}$ and proceeding as before,

$$I(f) = \int e^{\int b(s')\phi((s')\mu(ds')} \prod_s e^{-\phi^2(s)\mu(ds)} \sqrt{\frac{\mu(ds)}{\pi}}\, \delta\phi(s)$$

$$= \prod_s \left[\int e^{[b(s)\phi(s)-\phi^2(s)]/\sigma^2} \frac{1}{\sigma\sqrt{\pi}}\, \delta\phi(s)\right]$$

$$= \prod_s e^{\frac{1}{4} b^2(s)\mu(ds)} = e^{\sum_s \frac{1}{4} b^2(s)\mu(ds)}$$

$$I(f) = e^{\frac{1}{4}\int b^2(s)\mu(ds)}$$

Other examples of functionals which can be integrated explicitly will be given later on.

A rigorous derivation of these and other evaluations will be given in Appendices I, II, and III.

The question arises naturally whether the integral $I(\cdot)$ induces a measure in the space of functions $\{\phi\}$. That is indeed the case for certain sets of functions $\{\phi\}$. For example, the measure of the set of those functions ϕ such that

$$\lambda_1 \leq \int b(s)\phi(s)\mu(ds) \leq \lambda_2$$

is well-defined and can easily be computed. For the evaluation of this measure see Appendix I.

Another question which arises is whether or not it would have been possible to introduce this measure first and base the notion of integral on it. That can indeed be done provided the measure is totally additive, i.e., if the measure of the sum of infinitely many disjoint sets is the sum of the measures of the sets. The Wiener integral results from such an approach.

Whether or not the measure as described is totally additive depends on the manifold of functions adopted. In a way, it is very strange that up to now we have not been forced to specify this manifold except for the requirement that it include all piecewise constant functions and that all the functions in it be square-integrable.

Unless the manifold of functions $\{\phi(\cdot)\}$ is very specially chosen the measure will not be totally additive. For example the measure is not totally additive if the space of functions $\{\phi\}$ is taken to be the Hilbert space of those functions ϕ for which

$$\int_{-\infty}^{\infty} \phi^2(s)ds < \infty .$$

On the other hand, one may introduce the functions $\chi(\cdot)$ defined by the relation

$$\chi(s) = \int_0^s \phi(s')ds' \quad (0 \leq s \leq s_0)$$

where we have taken the ordinary Lebesgue measure differential ds for $\mu(ds)$ and have taken $\mathcal{S} = [0, s_0]$ for some $s_0 > 0$.

Furthermore, suppose we extend the space of functions $\{\chi\}$ to include all continuous functions. Then the measure induced by $I(\cdot)$ on the sets of those functions χ such that, for given λ_1, λ_2 and s,

$$\lambda_1 \leq \chi(s) \leq \lambda_2$$

is totally additive. This remarkable fact is part of Wiener's theory of integration. In fact, the integral I(f) just becomes the Wiener integral if it is expressed in terms of χ instead of ϕ and if the manifold of functions $\{\chi\}$ is taken to consist just of the continuous functions.*

In order to express the integral I(f) in terms of the Wiener integral I_W we introduce the functional h defined by the relation

$$h[\chi(\cdot)] = h[\int_0 \phi(s')ds'] = f[\phi(\cdot)].$$

Then I(f) is given in terms of the Wiener integral by

$$I(f) = I_W(h).$$

From our formal expression for I(f) we may derive a formal expression for $I_W(h)$. To this end we note that we may write symbolically

$$\delta\chi(s) = \delta\phi(s)\, ds \quad (s \neq 0)$$

and hence

$$\int_a \phi^2(s)ds = \int_y \frac{[\delta\chi(s)]^2}{ds}.$$

Accordingly we may write

$$I_W(h) = \int^{\infty} h(\chi)e^{-\int([\delta\chi(s)]^2/ds)} \prod_{s>0} \frac{1}{\sqrt{\pi ds}}\, \delta\chi(s)$$

Just as Wiener integral, the closely related integrals of Lessen and Kolmogoroff, which we shall discuss later, have this property of total additivity.

---

* C.f. Chapters IX and X.

The integrals of Wiener, Kolmogoroff and others were developed in connection with problems in probability. In this field total additivity is essential. In other connections, however, other properties may be more important than total additivity. In setting up a definition of an integral having these other properties it may be necessary to sacrifice total additivity.

I shall discuss one such other property which is important in connection with problems in the quantum theory of fields.

The functions ϕ may be considered as representatives of elements Φ in a linear space. These elements may just as well be given another representation, such as one in terms of functions ψ, where ϕ and ψ are connected by a linear transformation L. That is, $\phi = L\psi$. This transformation may, for example, be a generalized Fourier transformation,

$$\phi(s) = \int_0^\infty e^{-st}\psi(t)\mu^*(dt)$$

involving a measure differential $\mu^*(dt)$.

We want to make the assumption that the relation

$$\int_{\mathcal{S}} \phi^2(s)\mu(ds) = \int_{\mathcal{S}^*} \psi^2(t)\mu^*(dt)$$

holds for all pairs of functions, ϕ and ψ, connected by $\phi = L\psi$ and belonging, respectively, to appropriate base spaces, $\mathcal{S}$ and $\mathcal{S}^*$. In the present context it would be an undue restriction to insist on a linear measure, $\mu^*(dt) = cdt$. In particular, one must admit discrete measures. This is necessary if for example the functions ϕ and ψ are related by a Fourier series transformation.

In the same way in which we have defined the integral $I(f) = I_\mu(f)$ of functionals $f[\phi]$, we may define the integral $I_{\mu^*}(g)$ of functionals $g(\psi)$.

To a functional $f[\cdot]$ one may, in particular, assign the functional $g[\cdot]$ defined by

$$g[\psi] = f[L\psi].$$

Then one naturally expects that the integrals of these functionals are equal. I.e., one expects that

$$I_\mu{}^*(g) = I_\mu(f).$$

We shall say that this identity expresses the rotational invariance of the integral $I(\cdot)$. It is one of the objectives of these lectures to show that this rotational invariance obtains provided it is properly formulated.

The difficulty of formulating this statement results mainly from the difficulty in giving a precise meaning to the relation $g[\psi] = f[L\psi]$. Suppose the functional $f[\cdot]$ is compatible with respect to projections associated with partitions of the s-axis - then the functional $g(\psi) = f(L\psi)$ will not, in general, be compatible with respect to partitions of the t-axis. It will be necessary to define it as an ideal element associated with a sequence of functionals of ψ which are compatible with partitions of the t-axis.

This difficulty can be overcome. About five years ago I had described a way in which the invariance theorem can be proved. In this proof a detour via expansion of functionals in a series of Hermite polynomials was employed. To be sure, this approach is not quite satisfactory.*

In these considerations the linear space of elements $\mathfrak{P}$, represented alternatively by ϕ or ψ (functions on $\mathcal{S}$ and $\mathcal{S}^*$ respectively), is assumed to be a Hilbert space. As a consequence the total additivity of the measures is lost. It may turn out that the restriction to Hilbert spaces can be eliminated. Here is an open question which should, and probably can, be answered.**

---

* A more satisfactory approach to the problem of rotational invariance has been developed recently by H.N.Shapiro and myself and a related approach (for which see Chapters VI and VII) has been developed by J. Schwartz.

** As will be shown in Chapter XII, the total additivity can in fact be retrieved by enlarging the Hilbert space in a proper manner.

I should like to mention a simple problem which can be solved once the rotation invariance of the integration of functionals is established. Suppose one wants to evaluate the integral

$$I=\int^{\infty} \exp.\ \left\{Q[\phi] - \int_{\mathcal{S}} \phi^2(s)\mu(ds)\right\} \prod_s \sqrt{\frac{\mu(ds)}{\pi}}\ \delta\phi(s)$$

where

$$Q[\phi] = \iint_{\mathcal{S}} b(s',s'')\phi(s')\phi(s'')\mu(ds')\mu(ds'')$$

is the general quadratic polynomial functional on functions on $\mathcal{S}$. Then the following procedure suggests itself: one introduces functions ψ by a linear orthogonal transformation $\phi = L\psi$ such that the exponent $Q[\cdot]$ assumes diagonal form:

$$Q[\phi] = Q[L\psi] = \int_{\mathcal{S}^*} k(t)\psi^2(t)\mu^*(dt)$$

or, if the space $\{\psi\}$ may be taken as a sequence space, $\psi = \{\psi_1,\psi_2,\dots\}$,

$$Q[\phi] = a\sum_n k_n\psi_n^2 ,$$

while at the same time

$$\|\phi\| = \|\psi\|$$

or

$$\int_{\mathcal{S}} \phi^2(s)\mu(ds) = a\sum_n \psi_n^2$$

The transformation L then furnishes a rotation of the coordinate system to the principal axes of the quadratic form Q. By the rotation invariance of the integral one obtains

$$I = \int_e^{\infty - a\sum_n (1-k_n)\psi_n^2} \prod_n \sqrt{\frac{a}{\pi}}\, d\psi_n .$$

This integral can be evaluated directly. Its value is simply

$$I = \prod_n (1-k_n)^{-\frac{1}{2}} = e^{-\frac{1}{2}\sum_n \log (1-k_n)}$$

assuming that the series in the exponent converges absolutely. For a rigorous derivation of this formula see Appendix III.

Such evaluations play a considerable role in the theory of Brownian motion and related problems of statistical mechanics.

It should be mentioned* that, in addition to the Caussian measures, there exist other compatible measure differentials dm. Such measures have been proposed and investigated in the theory of probability.

As a somewhat freakish example we may give one of them by writing down the natural symbolic expression for the associated integral -

$$I(f) = \int^{\infty} f[\phi] \prod_s \frac{(ads)^{ds}}{\Gamma(ds)} \phi(s)^{(ds-1)} e^{-a\phi(s)ds}\, \delta\phi(s)$$

where it is assumed that $\phi(s) > 0$.

We mention in somewhat greater detail another, more reasonable, possibility - the differentials $dm_s(\phi)$ connected with the Poisson distribution in place of the Gaussian distribution.

We denote the independent functions by n(s) rather than $\phi(s)$. They are, in this case, supposed to be non-negative and integer-valued. We make no assumptions of integrability with respect to any measure $d\mu(s)$ and must accordingly employ a form of projection other than that in the foregoing.

We assume that, in any bounded interval $\mathcal{I}_\nu$, the function n(s) vanishes at all but a finite number of points. Then the projections to be used are constructed as follows:

Let $\{s_\nu\}$ ($\nu=0,\ldots r$) be a finite set of real numbers with $s_{\nu-1} < s_\nu$ and let $\mathcal{I}_\nu = [s_{\nu-1}, s_\nu)$. For $\nu=1,\ldots r$ choose a point c_ν in $\mathcal{I}_\nu$. Then define a projection P by

---

* and indeed, the matter will be covered at somewhat greater length in a later chapter.

$$[Pn](s') = \begin{cases} \sum_{s \in \vartheta_\nu} n(s) = n_\nu & (\text{if } s' = c_\nu, \nu = 1, \ldots r) \\ 0 & (\text{otherwise}). \end{cases}$$

This is again a function of the same type as n(s) and, clearly, if P' corresponds to a refinement of the partition giving rise to P as above, then PP'n = Pn.

With the aid of a measure differential $\mu(ds)$ we introduce differentials $d\bar{m}_\nu(z)$ determined by the relation

$$\int f(\Delta_\nu) d\bar{m}_\nu(\Delta_\nu) = \sum_{k=0}^{\infty} f(k) \frac{\Delta_\nu}{k!} e^{-\Delta_\nu}$$

which is to hold for arbitrary functions where

$$\Delta_\nu = \int_{\vartheta_\nu} \mu(ds).$$

We then define the differentials $dm_s(n(s))$ by

$$dm_s(n(s)) = d\bar{m}_\nu(n_\nu).$$

It is easily verified that these differentials $\{dm_s\}$ satisfy the compatibility condition introduced earlier in this chapter.

Employing these $\{dm_s\}$ we define the integral of a cylinder functional f[n] = f[Pn] as

$$I(f) = \sum_{n_1=0}^{\infty} \cdots \sum_{n_r=0}^{\infty} f_p(n_1, \ldots n_r) \prod_{\nu=1} \frac{(\Delta_\nu)^{n_\nu}}{n_\nu!} e^{-\Delta_\nu}$$

Let f[·] be the ideal functional associated with the Cauchy sequence of cylinder functionals $f_j[\cdot]$. For the limit as $j \longrightarrow \infty$ of the expression for $I(f_j)$ we then employ the symbol

$$I(f) = \sum_{n(\cdot)} f[n] \prod_s \frac{\mu(ds)^{n(s)}}{n(s)!} e^{-\mu(ds)} .$$

In case $\int \mu(ds) < \infty$ the factor $e^{-\int \mu(ds)}$ may be taken out. In this case the functional,

$$f_o[n] = \begin{cases} 1 & n(s) \equiv 0 \\ 0 & \text{otherwise} \end{cases}$$

is integrable. Evidently

$$I(f_o) = e^{-\int \mu(ds)}.$$

In case $\int \mu(ds) = \infty$ we have $I(f_o) = 0$; i.e., f_o is equivalent to the functional identically zero.

It is interesting to note that an integral

$$I'(f) = \sum_{n(s)} f[n] \prod_s \frac{[d\mu(s)]^{n(s)}}{n(s)!}$$

could also be defined by a procedure deviating slightly from that so far described. With respect to this new integral we always have

$$I'(f_o) = 1$$

independently of the character of $\int \mu(ds)$.

These facts are significant in the quantum theory of fields where f_o may be interpreted as the probability amplitude for the vacuum state.

Appendix I

In this appendix we shall establish, by formal manipulation, a formula for the integral of a function of a linear functional. Using this formula we will be able to evaluate the measures of certain sets as the integrals of their characteristic functionals.

Let $f[\phi] = h(L[\phi])$ where

$$L[\phi] = \int b(s)\phi(s)\mu(ds).$$

Let $\ell = \int b^2(s)\mu(ds)$. We shall show that

$$(1) \qquad I(f) = \frac{1}{\sqrt{\ell\pi}} \int_{-\infty}^{\infty} e^{-y^2/\ell}\, h(y)dy.$$

Let us write down the Fourier integral identity for h,

$$(2) \qquad h(x) = \frac{1}{2\pi} \iint e^{i(x-y)z} h(y)dydz,$$

and introduce the family of functionals

$$g_z[\phi]\, e^{izL[\phi]}$$

whose integrals, by the evaluation on page 13 of the text, are

$$I(g_z) = e^{-\frac{1}{4}\ell z^2}.$$

Putting $x = L[\phi]$ in (2) we get

$$f[\phi] = h(x) = \frac{1}{2\pi} \iint e^{izx} e^{-iyz}\, h(y)dydz$$

$$= \frac{1}{2\pi} \iint g_z[\phi] e^{-iyz} h(y)dydz.$$

In view of the linearity of $I(\cdot)$ we have, therefore,

$$I(f) = \frac{1}{2\pi}\iint I(gz)e^{-iyz}h(y)dydz$$

$$= \frac{1}{2\pi}\iint e^{-[\frac{1}{4}\ell z^2 - iyz]}h(y)dydz$$

$$= \frac{1}{\sqrt{\pi}}\int e^{-y^2/\ell}\ h(y)dy$$

which is formula (1).

Consider, now, the set S of those functions ϕ for which

$$\lambda_1 \leq \int b(s)\phi(s)\,\mu(ds) \leq \lambda_2$$

The measure of S may be regarded as the integral of its characteristic function which is

$$\chi_S[\phi] = \begin{cases} 1 & \phi\varepsilon S \quad \text{i.e. for } \lambda_1 \leq L[\phi] \leq \lambda_2 \\ 0 & \phi\varepsilon S \quad \text{i.e. for } \lambda_1 > L[\phi] \text{ or } L[\phi] > \lambda_2 . \end{cases}$$

This description may be written as $\chi_S[\phi] = h(L[\phi])$ where

$$h(x) = \begin{cases} 1 & \lambda_1 \leq x \leq \lambda_2 \\ 0 & \lambda_1 > x \text{ or } x > \lambda_2 \end{cases}$$

The integral of $\chi_S(\phi)$ — which is the measure of S — may then be evaluated by the formula (1). Thus we obtain

$$I(\chi_S) = \frac{1}{\sqrt{\ell\pi}}\int_{-\infty}^{\infty} e^{-x^2/\ell}\,h(x)dx = \frac{1}{\sqrt{\ell\pi}}\int_{\lambda_1}^{\lambda_2} e^{-x^2/\ell}\,dx$$

and so

$$\text{meas. } [S] = \frac{1}{\sqrt{\ell\pi}}\int_{\lambda_1}^{\lambda_2} e^{-x^2/\ell}\ dx.$$

This result may be expressed in terms of the standard error function

$$\psi[\xi] = \frac{1}{\sqrt{\pi}} \int_0^{\xi} e^{-x^2}\, dx$$

as

$$\text{meas. } [S] = \psi\left[\frac{\lambda_2}{\sqrt{t}}\right] - \psi\left[\frac{\lambda_1}{\sqrt{t}}\right]$$

Appendix II

In the text we evaluated — formally — the respective integrals of the functionals

$$f[\phi] = (L[\phi])^2 \text{ and } g[\phi] = e^{L[\phi]}$$

where

$$L[\phi] = \int b(s)\phi(s)ds \text{ for some } b(\cdot) \in \mathcal{L}_2.$$

In this appendix we wish to demonstrate the very existence of these integrals. We must show, then, that f and g may be represented as the limit of sequences of cylinder functionals Cauchy convergent in the metric induced by our integral.

As we have assumed b to be in $\mathcal{L}_2$ we may find a sequence $\{b_\nu\}$ of piecewise constant functions approximating b in the $\mathcal{L}_2$ sense so that

$$\int |b_\nu(s) - b(s)|^2 ds \to 0 \quad \text{as } \nu \to \infty .$$

Hence

$$\int |b_\nu(s) - b_\mu(s)|^2 ds \to 0 \quad \text{as } \nu,\mu \to \infty . \tag{1}$$

It follows that

$$\int |b_\nu^2(s)ds \to \int b^2(s)ds \quad \text{as } \nu \to \infty$$

and, therefore, that there is a constant $C > 0$ such that

$$\int b_\nu^2(s)ds < C \qquad (\nu = 1,2,\ldots). \tag{2}$$

We define the linear functionals $\{L_\nu\}$ by

$$L_\nu[\phi] = \int b_\nu(s)\phi(s)ds \qquad (\nu=1,2\ldots).$$

These functionals - and, of course, all functions of them - are cylinder functionals.

Let us consider first the case of $f[\phi] = L^2[\phi]$. We shall set

$$f_\nu[\phi] = L^2_\nu[\phi]$$

and show that

$$(+) \qquad I(|f_\nu - f_\mu|) \to 0 \qquad \text{as } \nu,\mu \to \infty .$$

Once this has been done it follows that there is an ideal functional $f[\cdot]$ corresponding to the "limit" of the sequence $f_\nu[\cdot]$. As it is well known, relation (+) implies, for almost all functions ϕ,

$$f_\nu[\phi] \to f[\phi] \qquad \text{as } \nu \to \infty .$$

The last relation justifies the identification of f with $\hat{f}$ and of $I(f)$ with $I(\hat{f})$, and thus proves the integrability of f. The evaluation of $I(\hat{f})$ follows by noting that for the cylinder functionals f_ν, the formal manipulations in the text are justifiable operations for finite dimensional integration; hence

$I(f_\nu) = \int b_\nu^2(s)ds$ and

$$I(\hat{f}) = \lim_\nu I(f_\nu) = \lim_\nu \int b_\nu^2(s)ds = \int b^2(s)ds$$

that is

$$I(f) = \int b^2(s)ds .$$

It only remains to show that

$$I(|f_\nu - f_\mu|) \to 0 \qquad \text{as } \nu,\mu \to \infty .$$

We shall use the inequality

$$(*) \quad |a^2-b^2| \le a(a^2+b^2) + 2a^{-1}(a-b)^2 \qquad \text{for } 0 < a \le 2$$

which is valid for any a and b.

For any given $\varepsilon > 0$ let us take $a = \varepsilon/3C$, where C is the constant appearing in (2). We assume ε to be so small that $a \leq 2$. By (1) there is an N such that, for $n,m > N$,

$$(3) \qquad \int |b_n(s) - b_m(s)|^2 ds < \tfrac{1}{2} a^2 C$$

Then, for any ϕ, we substitute

$$a = L_n[\phi] \text{ and } b = L_m[\phi]$$

into (*) getting, for each ϕ,

$$|f_n[\phi]-f_m[\phi]| = |L_n^2 - L_m^2| \leq aL_n^2 + aL_m^2 + 2a^{-1}(L_n-L_m)^2$$

and hence

$$I(|f_n-f_m|) \leq aI(L_n^2)+aI(L_m^2)+2a^{-1}I((L_n-L_m)^2)$$

As L_n^2, L_m^2 and $(L_n-L_m)^2$ are cylinder functionals, we can already evaluate their integrals - $I([\int a(s)\phi(s)ds]^2) = \int a^2(s)ds$ — and

$$I(|f_n-f_m|) \leq a\int b_n^2(s)ds+a\int b_m^2(s)ds+2a^{-1}\int |b_n(s)-b_m(s)|^2 ds$$

$$< 2aC +2a^{-1}\cdot \tfrac{1}{2} a^2 C = 3Ca = \varepsilon$$

Thus we have shown that

$$I(f_\nu - f_\mu|) \longrightarrow 0 \quad \text{as } \nu,\mu \longrightarrow \infty \ .$$

Next we consider the case of $g[\phi] = e^{L[\phi]}$. We let $g_\nu = e^{L_\nu}$ and make use of the inequality

$$(**) \quad |e^a-e^b| \leq a(e^{2a}+e^{2b}) + 4a^{-1}|a-b|^2 \qquad \text{for } a > 0,$$

valid for any a and b, to show that

$$I(g_\nu - g_\mu|) \longrightarrow 0 \quad \text{as } \nu,\mu \longrightarrow \infty \ .$$

The integrability of g and the evaluation of I(g) follow by the same arguments that were used in the discussion of f.

Let $D = e^{\frac{1}{2}C}$. Then, from (2), we get

$$(2') \qquad e^{\frac{1}{2}\int b_\nu^2(s)ds} < D .$$

For any $\varepsilon > 0$ let $a = \varepsilon/3D$. By (1) there is an $N' > 0$ such that, for all $n,m > N'$,

$$(3') \qquad \int |b_n(s) - b_m(s)|^2 ds < \frac{1}{4} a^2 D .$$

For any ϕ we substitute

$$a = L_n[\phi] \text{ and } b = L_m[\phi]$$

into (**) getting for each ϕ

$$|g_n[\phi] - g_m[\phi]| = |e^{L_n} - e^{L_m}| \leq ae^{2L_n} + ae^{2L_m} + 4a^{-1}(L_n - L_m)^2$$

and

$$I(|g_n - g_m|) \leq aI(e^{2L_m}) + 4a^{-1}I(L_n - L_m)^2) ,$$

Again we note that we can already evaluate $I(e^{L_n})$ and $I(e^{L_m})$ since L_n and L_m are cylinder functionals -- in fact,

$$I(e^{\int a(s)\phi(s)ds}) = e^{\frac{1}{4}\int a^2(s)ds}$$

and, therefore, using (2') and (3'),

$$I(|g_n - g_m|) \leq ae^{\frac{1}{2}\int b_n^2(s)ds} + ae^{\frac{1}{2}\int b_m^2(s)ds} + 4a^{-1}\int |b_n(s) - b_m(s)|^2 ds$$

$$< 2aD + 4a^{-1} \cdot \frac{1}{4} a^2 D = 3Da = \varepsilon ,$$

Hence,

$$I(|g_n - g_m|) \longrightarrow 0 \quad \text{as } n,m \longrightarrow \infty .$$

Thus the proof of the integrability of g is concluded. As before, in the case of f, we conclude

$$\lim_{n\to\infty} I(g_n) = \lim_{n} \left\{ e^{\frac{1}{2}\int b_n^2(s)ds} \right\}$$

and hence

$$I(g) = \left\{ e^{\frac{1}{2}\int b^2(s)ds} \right\}$$

Appendix III

Let $\mathcal{H}$ be a Hilbert space of elements, $\phi,\psi\ldots$, with the inner product devoted by (ϕ,ψ). Let K be a completely continuous operator acting on the elements ϕ of $\mathcal{H}$. Denote the eigen-values of K by $\{K_{\pm\sigma}\}$. We index these eigenvalues so that positive eigenvalues correspond to positive indices and negative eigenvalues to negative indices; each eigenvalue being repeated as often as necessary to match its multiplicity. We write

"$\sum_{\sigma}$" for "$\sum_{\sigma=1}^{n^+}$" where n^+ $(0 \le n^+ \le \infty)$ is the number of positive eigenvalues with their respective multiplicities and similarly for $\sum_{-\sigma}$.

We introduce the positive and negative traces of K —

$$TrK_+ = \sum_{\sigma} K_{+\sigma} \quad \text{and} \quad TrK_- = \sum_{-\sigma} K_{-\sigma} ,$$

and require that these each be finite —

$$TrK_+ < \infty \qquad \text{and} \quad TrK_- < \infty .$$

The trace of K is then given by their sum —

$$TrK = TrK_+ + TrK_- .$$

The difference of the positive and traces will be called the "absolute trace" —

$$Tr|K| = TrK_+ - TrK_-$$

We assume further that the eigenvalues have a bound less than 1,

$$K_\sigma \leq K < 1 .$$

These assumptions being made an operator $\log(I-K)$ can be defined. It is well known that this operator also possesses finite positive and negative traces and

$$(*) \qquad \mathrm{Tr}[\log(I-K)]_{\pm} = \sum_{\pm\sigma} \log(1-K_{\pm\sigma})$$

and that the right hand side of (*) converges absolutely.

We want to prove the

Theorem: Under the assumptions on K made above the functional

$$f(\phi) = e^{(\phi, K\phi)}$$

is integrable and

$$I(f) = e^{-\mathrm{Tr}[\log(I-K)]}$$

We shall first prove the

Lemma: Let $\{P_n\}$ be any sequence of projections, each with finite dimensional range, which strongly approach the identity. I.e., for all $\phi \in \mathfrak{H}$,

$$P_n\phi \longrightarrow \phi \qquad \text{as} \quad n \longrightarrow \infty .$$

Then, as $n \longrightarrow \infty$,

$$\mathrm{Tr}[\log(I-K^{(n)})] \longrightarrow \mathrm{Tr}[\log(I-K)]$$

where the operator $K^{(n)}$ is defined by

$$K^{(n)} = P_n K P_n .$$

Proof of the Lemma: From the maximum property of eigenvalues (see Courant-Hilbert, Volume I) it follows that

(a) $$0 \le \kappa_{\sigma}^{(n)} \le \kappa_{\sigma} \text{ and } 0 \ge \kappa_{-\sigma}^{(n)} \ge \kappa_{-\sigma}$$

where the $\{\kappa_{\pm\sigma}^{(n)}\}$ are the eigenvalues of $P_n K P_n$ indexed to correspond to the $\{\kappa_{\pm\sigma}\}$.

Also, we state the relation

(b) $$\kappa_{\pm\sigma}^{(n)} \to \kappa_{\pm\sigma} \quad \text{as } n \to \infty .$$

though we shall not prove it here.

It follows from the absolute convergence of the sums on the right in (*) that, for any $\varepsilon > 0$, we can find a $\zeta = \zeta(\varepsilon)$ such that

$$0 \le \sum_{\sigma > \zeta} \log (1-\kappa_{+\sigma}) < \varepsilon/4$$

and

$$0 \le \sum_{\sigma > \zeta} \log (1-\kappa_{-\sigma}) < \varepsilon/4 .$$

Then we have, for each n

(c) $$0 \le \sum_{\sigma > \zeta} \log (1-\kappa_{\sigma}^{(n)}) - \sum_{\sigma > \zeta} \log (1-\kappa_{\sigma}) < \varepsilon/4$$

and also

(c') $$0 \le \sum_{\sigma > \zeta} \log (1-\kappa_{-\sigma}) - \sum_{\sigma > \zeta} \log (1-\kappa_{-\sigma}^{(n)}) < \varepsilon/4$$

Further, by (b), we may choose $N = N(\varepsilon,\zeta)$ so that, for every $n > N$,

(d) $$0 \le \sum_{\sigma \le \zeta} \log (1-\kappa_{\sigma}^{(n)}) - \sum_{\sigma \le \zeta} \log (1-\kappa_{\sigma}) < \varepsilon/4$$

and also

(d') $$0 \le \sum_{\sigma \le \zeta} \log (1-\kappa_{-\sigma}) - \sum_{\sigma \le \zeta} \log (1-\kappa_{-\sigma}^{(n)}) < \varepsilon/4.$$

Adding (c) and (d) we get

$$0 \leq \sum_{\sigma} \log\,(1-\kappa_{\sigma}^{(n)}) - \sum_{\sigma} \log\,(1-\kappa_{\sigma}) < \varepsilon/2$$

or

$$0 \leq \mathrm{Tr}[\log(I-K^{(n)})]_{+} - \mathrm{Tr}[\log(I-K)]_{+} < \varepsilon/2.$$

Similarly we obtain

(e') $$0 \leq \mathrm{Tr}[\log\,(I-K^{(n)})]_{-} - \mathrm{Tr}[\log(I-K)]_{-} < \varepsilon/2.$$

Relations (e) and (e') evidently yield the statement of the Lemma.

Having this we go on to the

Proof of the Theorem:

We wish to show that I(f) exists and

$$I(f) = e^{-\mathrm{Tr}[\log(I-K)]}$$

We consider the approximating cylinder functionals f_n given by

$$f_n(\phi) = f(P_n\phi) = e^{(P_n\phi, KP_n\phi)} = e^{(\phi, P_nKP_n\phi)}.$$

Clearly

$$I(f_n) = \int^{\infty} e^{-(\phi,(I-K^{(n)})\phi)} \prod_s \sqrt{\frac{ds}{\pi}}\,\delta\phi(s) = e^{-\mathrm{Tr}[\log(I-K^{(n)})]}$$

as may be observed by evaluation of the integral which is here finite. It follows from the Lemma (see the second part of Appendix I) that

$$I(|f_n - f_m|) \longrightarrow 0 \quad \text{as } n,m \longrightarrow \infty.$$

We are, therefore, permitted to identify f with the ideal functional $\lim f_n$ getting $I(f) = e^{-\mathrm{Tr}[\log(I-K)]}$ as the limit of $I(f_n)$ by the Lemma.

Chapter II

Remarks on the Quantum Theory of Fields

I should like to make some remarks about the role of the integration of functionals in Quantum Theory.

In the quantum theory of particles the "state" of a system is described by a "wave function", i.e., a function $F(\phi_1 \ldots)$ of the possible values of a complete set of commuting "observables" $[\phi]$, which for example in the case of a single particle, might be the three coordinates of position. In the quantum theory of fields a function of infinitely many commuting observables is needed for a complete description of the state of the field. One might take as a set of observables the values of a potential ϕ at all points x $(= \{x_1, x_2, x_3\})$ of space in which case a wave function would be a function of all these values $\phi(x)$; in other words, a wave function would then be a functional $F[\phi]$ of the function $\phi(x)$.

In the quantum theory of particles one introduces the concept of the probability $Pr(\mathcal{S})$ that measurement of the observables yields values belonging to the set $\mathcal{S}$. For a single particle $\mathcal{S}$ would be a set of ordered triples (i.e. a point-set in 3-space) and $Pr(\mathcal{S})$ would be the probability that the position, $\phi = \{\phi_1, \phi_2, \phi_3\}$, of the particle lies in the point-set $\mathcal{S}$. This probability is given by the integral,

$$Pr(\mathcal{S}) = \iiint_{\mathcal{S}} |F(\phi_1, \phi_2, \phi_3)|^2 \, d\phi_1 d\phi_2 d\phi_3 ,$$

Where $F = F(\phi_1, \phi_2, \phi_3)$ is the wave function associated with observation of position.

In field quantum theory also one should be able to introduce the probability, $Pr(\mathcal{S})$, that measurement of the set of values $\phi(x)$ of the potential ϕ at each point, $x = \{x_1, x_2, x_3\}$, in space yields a function $\phi(\cdot)$ which belongs to a certain set $\mathcal{S}$ of such functions. Naturally one expects that $Pr(\mathcal{S})$ is given here by an analogous "integral" of the square of (the absolute value of) the wave functional $F[\phi]$ associated with observation of potential.

The question arises; how to define such an integral?

In describing the proper form of this integral we shall assume that the integration is extended over the whole manifold of potential functions $\{\phi\}$. This is not really restrictive as, for any set $\mathcal{S}$ of an appropriate type, integration over $\mathcal{S}$ may be defined by replacing the functional F by a functional $F_{\mathcal{S}}$ equal to F on $\mathcal{S}$ and vanishing elsewhere.

The kind of integral needed to express these probabilities depends on the kind of field to be investigated. Let us assume that the field is "free" (i.e., not under the influence of external charges) and, further, that it is in a "vacuum state". This is the state in which the energy assumes its lowest value. For reasons which we shall give later such a vacuum state is characterized by the wave functional,

$$F[\phi] = ce^{-\frac{1}{2}\int\int[\gamma\phi(x_1,x_2,x_3)]^2 dx_1 dx_2 dx_3} = cF_0[\phi]$$

where γ is the operator such that, for all ϕ, $\gamma^4\phi = -\Delta\phi$

(Δ is the Laplacian operator ; $\Delta\phi = \frac{\partial^2\phi}{\partial x_2} + \frac{\partial^2\phi}{\partial x_2^2} + \frac{\partial^2\phi}{\partial x_3^2}$)

and c is a normalizing constant so chosen that the integral of $|F[\phi]|^2$ - in the sense to be defined - has the value 1.

It is convenient to consider and introduce, instead of $F[\phi]$ the functional $f[\phi]$ by the formula

$$F[\phi] = e^{-\frac{1}{2}\int[\gamma\phi(x)]^2 dx} \qquad f[\phi] = f[\phi]F_0[\phi] .$$

Then the integral in question is given by the expression,

$$I_\gamma(|f|^2) = \int^{\infty} |f[\phi]|^2 e^{-\int[\gamma\phi(x)]^2 dx} \prod_x \sqrt{\frac{dx}{\pi}}\, \delta\gamma\phi(x)$$

which differs from the integral which we have discussed earlier in that $\gamma\phi(x)$ takes the place of $\phi(x)$.

Note that the vacuum state is now described by

$$f[\phi] \equiv \text{const.} = c \ .$$

Note, also that the completion of the manifold of functionals f for which $I_\gamma(|f|^2) < \infty$ forms a Hilbert space, $\mathcal{H}$, with a scalar product induced by the norm, $||f|| = [I_\gamma(|f|^2)]^{1/2}$.

The fact that the integral I_γ is written in terms of $\gamma\phi$ instead of ϕ is not quite satisfactory. One should really like to write the integral as an integral over the space of functions ϕ rather than $\gamma\phi$. It is true that the Jacobian of the transformation γ is infinite - as is easily verified. Still, it might be possible to write this determinant - symbolically - in the form $\prod_x j(x,dx)$, and to find differentials, essentially given by

$$m[\delta\phi(x)] = j(x,dx)\sqrt{\frac{dx}{\pi}}\,\delta\phi(x),$$

which would be compatible with partitions.

A more satisfactory form of the integral is obtained through the use of the Fourier transformation on the potential functions. For this the rotational invariance of the integral $I(\cdot)$ is necessary. Although this could be done we shall not describe this transformation.

The wave functions, $F[\phi] = f[\phi]F_0[\phi]$ that characterize possible states of the field are those in $\mathcal{H}$ i.e. those for which $I_\gamma(|f|^2) < \infty$. Among these states, one of particular interest is that which would be the vacuum state if electric charges were distributed over the field. Let $\hat{\phi} = \hat{\phi}(x)$ be the Coulomb potential which would result from such charges. The theory then yields that the wave function of the "modified vacuum state" is given by

$$f[\phi] = c\hat{f}(\phi) = ce^{\int[\gamma\phi(x)][\gamma\hat{\phi}(x)]dx}$$

where the value of the normalizing constant c is determined from the condition $I_\gamma(|f|^2) = 1$. With the aid of an evaluation given before (see p.I-13) we find

$$I_\gamma(|\hat{f}|^2) = |c|^2 e^2 \Gamma \text{ with } \Gamma = \frac{1}{2} \int [\gamma\hat{\phi}(x)]^2 \, dx$$

and hence

$$|c| = e^{-\Gamma}$$

provided the integral Γ is finite.

Whether or not Γ is finite depends on the distribution of the charges. It is easily seen that Γ is infinite if the total charge - the integral of the charge density over all space - does not vanish. We then speak of non-balancing charges. Since this condition is equivalent to the non-vanishing of the Fourier transform of the potential ϕ for the frequency zero, one speaks of an "infra-red catastrophe" if the charges do not balance.

A striking feature of such a catastrophe is this: If the field could be in a vacuum state in the absence of charges it cannot be in a vacuum state in the presence of non-balancing charges and vice versa.

It is a remarkable fact that an integral - of course, different form $I_\gamma(|f|^2)$ can be introduced with respect to which the functional $|\hat{f}|^2$ is integrable - no matter whether or not it is catastrophic.

One need only take

$$\hat{I}_\gamma(f) = \int^{\infty} \left|\frac{f[\phi]}{\hat{f}[\phi]}\right|^2 e^{-\int [\gamma\phi(x) - \gamma\hat{\phi}(x)]^2 dx} \prod_x \sqrt{\frac{dx}{\pi}}\, \delta\gamma\phi(x).$$

This expression for an integral may be given meaning in the same way as $I(|f|^2)$ except that one must begin with functionals $f[\phi]$ for which

$$\tilde{f}[\tilde{\phi}] = \frac{f[\hat{\phi}+\tilde{\phi}]}{f[\hat{\phi}+\phi]} \text{ is a cylinder functional of } \tilde{\phi}.$$

In case $\Gamma < \infty$ the two integrals are related by the formula

$$\hat{I}_\gamma(|f|^2) = e^{-2\Gamma} I_\gamma(|f|^2)$$

and the corresponding Hilbert spaces, $\hat{\mathcal{H}}$ and $\mathcal{H}$, of square integrable functionals, are the same. If $\Gamma = \infty$, however, these Hilbert spaces are different.

The possibility of different types of Hilbert spaces of states - one containing a free field vacuum state while the other does not contain such a state - was observed independently by van Hove and myself several years ago in a somewhat different setting. Van Hove's arguments made it very likely that the difficulties of field quantum theory which involve the occurence of undesirable infinities are closely connected with this possibility.

In the present problem one may say that transition from one of these Hilbert spaces to the other can be accomplished through multiplication by a factor (which actually is infinite). This procedure appears to be related to - though different from - the renormalization procedure in field quantum theory.

Integration of the squares of wave functions is only a preliminary problem in the quantum theory of fields. A major problem which might afterwards be attacked after the problem of integration has been settled concerns the formulation of differential equations for wave functions. The solutions of such equations will yield the wave functions characterizing specific states - e.g., the vacuum state - under various circumstances.

It is, therefore, necessary to investigate differential operators which act on functionals. A functional is a function of the infinitely many variables $\phi(x)$, x being an index ranging over 3-space, and it is natural to attempt to introduce a concept of partial differentiation, $\delta/\delta\phi(x)$, with respect to one of these variables. Instead of $\delta/\delta\phi(x)$, we shall define a related operator denoted by $\delta/\delta\phi(x)dx$.

We first consider cylinder functionals,

$$f[\phi] = f_P(\phi_1 \ldots \phi_n),$$

with

$$\phi_\nu = \int_{\mathcal{C}_\nu} \phi(x')dx' / \Delta_\nu \qquad (\nu = 1 \ldots n)$$

where

$$\Delta_\nu = \int_{\mathcal{C}_\nu} dx' \qquad (\nu = 1 \ldots n).$$

Here we denote by $\mathcal{C}_\nu$ disjoint parallelopipeds called "cells". If f is a cylinder functional as above it will be independent of the value of $\phi(\cdot)$ at $\phi(x)$ any point x lying in none of these n cells; we set $\delta f[\phi]dx = 0$ for such x. If x lies in the cell $\mathcal{C}_\nu$, $(\nu=\nu(x))$, we set

$$\frac{\delta}{\delta\phi(x)dx} f[\phi] = \frac{1}{\Delta_\nu} \frac{\partial}{\partial \phi_\nu} f_P(\phi_1 \ldots \phi_n).$$

For an arbitrary function w(x), we then form the operator such that

$$\int w(x) \frac{\delta}{\delta\phi(x)dx} dx f[\phi] = \sum_{\nu=1}^{n} \int_{\mathcal{C}_\nu} w(x)dx \frac{1}{\Delta_\nu} \frac{\partial}{\partial \phi_\nu} f_P(\phi_1 \ldots \phi_n).$$

Employing the mean values

$$w_\nu = \int_{\mathcal{C}_\nu} w(x)dx / \Delta_\nu$$

and canceling the differential dx, we write symbolically

$$\int w(x) \frac{\delta}{\delta\phi(x)} = \sum_{\nu=1}^{n} w_\nu \frac{\partial}{\partial \phi_\nu}.$$

This operator, so far defined only for cylinder functionals, can be extended to a larger class of functionals by the process of closure: if the sequences, $\{f^\sigma\}$ of cylinder functionals and $\left\{\int w(x) \frac{\delta}{\delta\phi(x)} f^\sigma\right\}$, converge - with respect to the norm adopted - to limits, $f = f[\phi]$ and $g = g[\phi]$, respectively,

we set

$$\int w(x) \frac{\delta}{\delta\phi(x)} f = g.$$

E. g., for a quadratic functional:

$$f_2[\phi] = \iint b(x',x'')\phi(x')\phi(x'')dx'dx''$$

which, in an obvious manner (see page I-2), can be approximated by cylinder functionals, we obtain the formula,

$$\int w(x) \frac{\delta}{\delta\phi(x)} f_2[\phi] = 2\iint b(x',x'')w(x')\phi(x'')dx'dx''$$

or, symbolically,

$$\frac{\delta}{\delta\phi(x)} f_2[\phi] = 2dx \int b(x,x')\phi(x')dx'$$

In a similar way we may define second derivatives,

$$\delta^2/\delta\phi(x')\delta\phi(x'')$$

and, in particular, the "Laplacian"

$$L = \int [\frac{\delta}{\delta\phi(x)dx}]^2 dx$$

or simply

$$L = \frac{\delta^2}{[\delta\phi(x)]^2 dx}$$

Again considering the example, $f_2[\phi]$, we obtain

$$Lf_2[\phi] = 2\int b(x,x)dx \ ;$$

i.e. twice the trace of the operator with kernel $b(x',x'')$.

The quantum theoretical description of a free field involves the Hamiltonian energy operator,

$$H = -L + \int [\gamma^2\phi(x)]^2 dx + E_o$$

where L is the Laplacian just introduced, γ is the operator such that $\gamma^4 = -\triangle$ (here $\triangle = D^2$), introduced earlier, and E_0, the zero energy, is a suitably chosen constant. The operator H is to be applied to wave functions $F[\phi]$. The vacuum state of the field is then characterized by a wave function $F_0[\phi]$ which is an eigenfunction of H corresponding to the lowest eigenvalue. The constant E_0 is determined by the condition that this least eigenvalue be zero.

Now the functional,

$$F_0[\phi] = e^{-\frac{1}{2}\int[\gamma\phi(x)]^2dx},$$

introduced earlier, is such an eigenfuction. In order to evaluate $HF_0[\phi]$ formally, we first derive

$$\frac{\delta}{\delta\phi(x)} F_0[\phi] = -\gamma^2\phi(x)dx\, F_0[\phi]$$

and then

$$L\, F_0[\phi] = \left\{\int[\gamma^2\phi(x)]^2dx - \int\gamma^2\right\} F_0[\phi] \quad ;$$

hence we obtain

$$HF_0[\phi] = \int\gamma^2 + E_0 \quad .$$

To give the expression $\int\gamma^2$ a meaning we may regard γ^2 as an integral operator with kernel $\gamma^2(x',x'')$. Then γ^2 is the trace of the operator:

$$\int\gamma^2 = \int\gamma^2(x,x)dx$$

The kernel $\gamma^2(x',x'')$ is singular like $|x'-x''|^{-4}$ for $x' = x''$. Therefore the trace $\int\gamma^2$ is infinite. The constant E_0 determined by

$$E_0 = -\int\gamma^2$$

is then infinite. Then it is seen that the operator H cannot be applied in a proper sense.

This difficulty can easily be overcome - as has been clear since the theory of fields was developed by Pauli and Heisenberg in 1929. One defines the operator H as applicable to wave functions $F[\phi]$ of the form.

$$F[\phi] = f[\phi]\, F_0[\phi] ,$$

by the formula,

$$HF[\phi] = \left\{-Lf[\phi] + 2\int \phi(x) \frac{\delta}{\delta\phi(x)} f[\phi]\right\} F_0[\phi]$$

$$= \left\{\hat{H} f[\phi]\right\} F_0[\phi]$$

which would result by carrying out the operation L formally. The class of functionals $f[\phi]$ can then be so chosen that the operator $\hat{H} = -L+2\int \phi(x) \frac{\delta}{\delta\phi(x)}$ are applicable in the sense explained before.

One of the fundamental requirements of quantum theory is that the operators corresponding to observables be Hermitian with respect to the unit form in terms of which the basic probabilities are expressed. Let us note that the operator $\hat{H}$ is Hermitian with respect to the unit form

$$I_\gamma(|f|^2) = \int^{\infty} |f[\phi]|^2 \, e^{-\int [\gamma\phi(x)]^2 dx} \prod_x \sqrt{\frac{dx}{\pi}}\, \delta\gamma\phi(x) ,$$

as is easily verified. This fact may also be expressed by saying that the operator H is Hermitian with respect to the unit form,

$$J_\gamma(|f|^2) = \int^{\infty} |F[\phi]|^2 \prod_x \sqrt{\frac{dx}{\pi}}\, \delta\gamma\phi(x).$$

It is for this reason that the expression $I_\gamma(|f|^2)$ - or $J_\gamma(|f|^2)$ - furnishes the proper integral with which to express the basic probabilities.

In conclusion I should like to mention a related problem in which one of the typical - as yet unsolved - difficulties of field quantum theory appears.

The problem is that of the proper definition and spectral resolution of a Hamiltonian operator which may be written formally as

$$H = -L + \int [\gamma^2 \phi(x)]^2 dx + c \int \phi^4(x) dx + E_0 .$$

This differs from the Hamiltonian previously described by a term of fourth degree.

Let us first suppose that the symbol γ stands for a constant. Further, let us replace the operator H by the operator,

$$H_p = -L + \gamma^2 \sum_{\nu=1}^{n} \phi_\nu^2 \Delta_\nu + c \sum^{n} \phi_\nu^4 \Delta_\nu + E_0 ,$$

which transforms cylinder functionals compatible with the partition $\mathcal{P}$ into such cylinder functionals. The spectral resolution of the operator H_p can be obtained by separation of variables and an examination of the spectrum of H_p shows that it approaches a definite limit upon refining the partition if one lets the coefficient c depend on the mesh width($\Delta_\nu = \Delta$, assumed independent of ν)in such a way that $\frac{c}{\Delta} = u$ remains constant as the partition is refined and Δ approaches zero. The limit operator should than symbolically be written in the form,

$$H = -L + \int [\gamma^2 \phi(x)]^2 dx + \gamma \int \phi^4(x) [dx]^2 + E_0 .$$

In the relation

$$c = \gamma dx,$$

we meet a genuine "renormalization". One may expect that a similar renormalization will be necessary for the operator H when γ is not constant but is equal to $\sqrt[4]{-D^2}$.

For a complete treatment of these operators it would be necessary to define the proper integral to be used as unit form. It is not obvious how to do this even in the case where γ is taken as a constant. If this problem were solved the notion of renormalization would have been given — in one case at least — a legitimate mathematical meaning.

Chapter III

Compatible Measures

§ 1. Introduction

At this point in the development of the seminar we are in the process of considering various precedures and ideas around which future discussions will center. We now wish to comment on a sidelight of the previous discussion, namely that of determining "compatible ϕ-measures" within the framework of integration of functionals proposed thus far. Only Gaussian measures have been used, and the question of whether or not other measures are suitable arises naturally. We shall see that there are a large number of other measures which can be used.

We consider $\underline{\Phi}$, a set of functions $\phi(s)$ defined over a certain interval I of the s-axis (possibly the whole s-axis). We assume that with every partition ρ of this interval there is associated a mapping P of $\underline{\Phi}$ into itself, which we call the <u>projection</u> associated with the partition.

<u>Definition</u>: The <u>projections</u> are said to be <u>compatible</u> if for any two partitions ρ, ρ', where ρ' is a refinement of ρ, the corresponding projections P, P' satisfy

$$PP' = P. \tag{1.1}$$

That is, for every $\phi \in \underline{\Phi}$, $PP'\phi = P\phi$.

We further assume that with each projection P there is associated a mapping of the pairs (I_ν, ϕ), (where I_ν are the subintervals of ρ), into the real numbers:

$$P(I_\nu, \phi) = \phi_\nu$$

which is such that $P\phi$ is uniquely determined by the $\{\phi_\nu\}$.

In what follows we restrict ourselves to partitions ρ which consist of only a finite number of intervals. In the case where I is the entire s-axis, there will be two infinite intervals in a partition.

In addition to assuming that partitions are compatible we shall assume the existence of a set function L(I) such that if $\{I_{\nu 1}, \ldots, I_{\nu K_\nu}\}$ is a partition of the interval I_ν we have

$$L(I_\nu)\,\phi_\nu = \sum_{\mu=1}^{K_\nu} L(I_{\nu\mu})\phi_{\nu\mu} \tag{1.2}$$

(wherein I_ν is an interval of the partition ρ, the $I_{\nu\mu}$ are intervals of a refinement ρ', and $\phi_{\nu\mu} = P'(I_{\nu\mu}, \phi)$.

By way of illustration, we note that the examples of earlier discussions fit into the above framework. If we let Φ be the set of all integrable functions on the interval I, we can define $P\phi$ for a partition ρ into subintervals I_ν, to be the step function which has the constant value

$$\phi_\nu = \frac{\int_{I_\nu} \phi\,ds}{\int_{I_\nu} ds}$$

in the interval I_ν; and take $P(I_{\nu,\phi}) = \phi_\nu$. If we set $L(I) = \int_I ds$ the condition (1.2) will be satisfied. Again, let Φ be the set of all functions which are zero except at a countable number of points in I, and for which the series of non-zero values converges absolutely. In this case define P_ϕ to be the function which is zero except at the end-point of each I_ν at which point it has the value

$$\phi_\nu = \sum_{s \in I_\nu} \phi(s) \quad ,$$

that is, the sum of the non-zero values of ϕ for $s \in I_\nu$. Taking $P(I_\nu, \phi) = \phi_\nu$, and $L(I) \equiv 1$, (1.2) is satisfied.

2. ϕ-measures, compatibility

Next we assume that to each interval I_ν in a partition ρ of I there is associated a distribution function $m_{I_\nu}(x)$, that

is, a non-decreasing function which is $\begin{cases}0\\1\end{cases}$ at $x = \begin{cases}-\infty\\+\infty\end{cases}$. We shall refer to these distributions as the ϕ-distributions and to their associated measures $dm_{I_\nu}(x)$ as the ϕ-measures.

Definition: A function $f[\cdot]$ defined over Φ is called a cylinder functional if there is a partition P such that for all $\phi \in \Phi$

$$f[P\phi] = f[\phi] \quad . \tag{2.1}$$

We note immediately that if P' is a refinement of P, and (2.1) holds, then using (1.1) we have

$$f[P'\phi] = f[PP'\phi] = f[P\phi] = f[\phi] \quad .$$

Thus, if a functional is a cylinder functional "with respect to a given partition" then it is also a cylinder functional with respect to any refinement of that partition.

If f is any functional defined over Φ, the functional $\hat{f}$ defined by

$$\hat{f}(\phi) = f(P\phi)$$

(where P is the projection of a partition P) is a cylinder functional. This follows from the fact that $PP\phi = P\phi$ and

$$\hat{f}[P\phi] = f[PP\phi] = f[P\phi] = \hat{f}[\phi] \quad .$$

Definition: If f is a cylinder functional with respect to P define its integral over Φ to be

$$\int^K f(P\phi)dm_{I_1}(\phi_1)dm_{I_2}(\phi_2) \ldots dm_{I_K}(\phi_K) \tag{2.2}$$

where the I_ν ($\nu=1,\ldots,K$) are the intervals of the partition P.

Here and elsewhere, except where the contrary is specified, all ordinary integrals are to be taken from $-\infty$ to $+\infty$. Note also that in (2.2), since $P\phi$ is uniquely determined by $\phi_1,\ldots,\phi_K$ we consider $f(P\phi)$ as a function of K variables, i.e., $f(P\phi) = f_P(\phi_1,\ldots,\phi_K)$ in the obvious sense.

Definition: The ϕ-measures are said to be compatible if and only if for every functional $f[\cdot]$ which is a cylinder functional with respect to a partition ρ' we have, for every refinement ρ' of ρ,

$$\int f_{\rho}(..\phi_{\nu}..) \prod_{\nu} dm_{I_{\nu}}(\phi_{\nu}) = \int f_{\rho'}(..\phi_{\mu\nu}..) \prod_{\mu,\nu} dm_{I_{\mu\nu}}(\phi_{\mu\nu})$$

(where the $I_{\mu\nu}$ are the intervals of the partition ρ').

If a functional f is a cylinder functional with respect to two partitions ρ_1 and ρ_2, it is also a cylinder functional with respect to their common refinement ρ_{12}. Thus, compatibility of the ϕ-measures is equivalent to the assertion that (2.2) defines the integral of a cylinder functional independently of the choice of the partition ρ such that $f[P\phi] = f[\phi]$.

§ 3. Necessary and sufficient conditions for compatibility of ϕ-measures.

Within the framework set down in the previous sections we now consider the problem of establishing conditions which will insure the compatibility of the ϕ-measures. In this connection our main result is:

Theorem 3.1 Assuming that the projections and the set function L(I) satisfy (1.2) a necessary and sufficient condition for the compatibility of the ϕ measures is that for the decomposition of any interval I_0 into two disjoint subintervals I_1 and I_2, we have

$$(3.1) \qquad m_{I_0}\left(\frac{\omega}{L(I_0)}\right) = m_{I_1}\left(\frac{\omega}{L(I_1)}\right) * m_{I_2}\left(\frac{\omega}{L(I_2)}\right) ,$$

where the right side of (3.1) is the convolution of the two distributions indicated.

The convolution of two functions F and G is defined by

$$F(w) * G(w) = \int_{-\infty}^{\infty} F(w-t)\, dG(t) .$$

It can be easily checked that for distribution functions this operation is both commutative and associative.

Proof: Necessity: Assume the ϕ-measures compatible. Consider the typical case where I_0 is an interval of the partition P and the refinement P' is obtained from P by dividing I_0 into two disjoint subintervals I_1 and I_2. We must, then, have

$$(3.2)\qquad \int_{(\phi_0)} f_P(\phi_0)\,dm_{I_0}(\phi_0) = \int_{(\phi_2)}\int_{(\phi_1)} (\phi_1,\phi_2)\,dm_{I_1}(\phi_1)\,dm_{I_2}(\phi_2)\ ,$$

(where $\phi_0 = P(I_0,\phi)$, $\phi_1 = P'(I_1,\phi)$, $\phi_2 = P'(I_2,\phi)$) for any functional $f[\cdot]$ which is a cylinder functional with respect to P and such that $f[P\phi] = f_P(\phi_0)$. Letting $L_0 = L(I_0)$, $L_1 = L(I_1)$, and $L_2 = L(I_2)$ we obtain from (1.2)

$$\int_{(\phi_2)}\int_{(\phi_1)} f_{P'}(\phi_1,\phi_2)\,dm_{I_1}(\phi_1)\,dm_{I_2}(\phi_2)$$

$$= \int_{(\phi_2)}\int_{(\phi_1)} f_P\left(\frac{L_1\phi_1+L_2\phi_2}{L_0}\right)\,dm_{I_1}(\phi_1)\,dm_{I_2}(\phi_2)$$

$$= \int_{(\phi_2)}\int_{(\phi_0)} f(\phi_0)\,d_{(\phi_0)}m_{I_1}\left(\frac{L_0\phi_0-L_2\phi_2}{L_1}\right)dm_{I_2}(\phi_2)$$

$$= \int_{(\phi_0)} f_P(\phi_0)\,d_{(\phi_0)}\left[\int_{(\phi_2)} m_{I_1}\left(\frac{L_0\phi_0-L_2\phi_2}{L_1}\right)dm_{I_2}(\phi_2)\right].$$

Thus we obtain

$$(3.3)\qquad \int_{(\phi_0)} f_P(\phi_0)\,dm_{I_0}(\phi_0) = \int_{(\phi_0)} f_P(\phi_0)\,d_{(\phi_0)}\left[m_{I_1}\left(\frac{L_0 w}{L_1}\right) * m_{I_2}\left(\frac{L_0 w}{L_2}\right)\right]_{w=\phi_0}$$

Since in (3.2) we may take for $f_P(\phi_0)$ the characteristics of any interval, it follows that

$$(3.4) \qquad m_{I_0}\left(\frac{w}{L_0}\right) = m_{I_1}\left(\frac{w}{L_1}\right) * m_{I_2}\left(\frac{w}{L_2}\right)$$

Sufficiency: Assuming (3.4), we may then assert (3.3) and reversing the steps of the necessity proof obtain (3.2) for any integrable functional such that $f[P\phi] = f[\phi]$. This in turn yields easily the compatibility of the ϕ-measures.

By a slight modification of the argument used in proving the above one can also prove:

Theorem 3.2 Assuming that the ϕ-distributions satisfy (3.1), the ϕ-measures are compatible if and only if (1.2) is satisfied (except possibly on a set of $\prod_\nu dm_{I_\nu}(\phi_\nu)$-measure zero).

§ 4. The construction of compatible ϕ-measures

We proceed next to note some simple sequences of Theorem 3.1. Letting $a(I)$ denote the mean of $m_I(w)$, and $\sigma^2(I)$ the variance of $m_I(w)$, defined by

$$a(I) = \int x dm_I(x) \quad , \qquad \sigma^2(I) = \int [x-a(I)]^2 \, dm_I(x) \quad ,$$

we obtain from (3.2) the relations

$$(4.1) \qquad L(I_0)a(I_0) = L(I_1)a(I_1) + L(I_2)a(I_2)$$

$$(4.2) \qquad L^2(I_0)\sigma^2(I_0) = L^2(I_1)\sigma^2(I_1) + L^2(I_2)\sigma^2(I_2)$$

(4.1) asserts simply that $L(I)a(I)$ must be an additive set function, which we denote by $A(I)$, i.e.

$$(4.3) \qquad A(I) = L(I)a(I)$$

On the other hand, (4.2) requires that $[L(I)\sigma(I)]^2$ be a non-negative additive set function, which we denote by $S(I)$. I.e.,

$$(4.4) \qquad S(I) = [L(I)\sigma(I)]^2 \, .$$

Thus in attempting to construct compatible ϕ-measures we are led to consider any stable distribution $F(x)$, i.e., one

such that the family $\{F(ax+b)\}$ $(a > 0)$ is closed under convolution. Then we pick any additive set function $A(I)$, any non-negative one $S(I)$, and any $L(I) \neq 0$. Then we can solve (4.3) and (4.4) for $a(I)$ and $\sigma^2(I)$, and arrange for (3.2) to hold. Finally, and this would be the critical step, one should then ascertain whether or not a compatible system of projections exist which satisfy (1.2). In any event one might view this as a restriction on the function $L(I)$ and take only those $L(I)$ for which this can be done.

We next proceed to indicate what the above construction would yield when applied to the most familiar examples of stable distributions.

Example I: The Gaussian Distribution

The standard distribution function

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}}\, du$$

which has mean 0 and variance 1, generates the stable family consisting of the Gaussian distribution functions

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{u-a}{\sigma}\right)^2}\, du$$

with mean a and variance σ^2.

Thus, if we take as our ϕ-measures

$$dm_I(\phi) = \frac{L(I)}{\sqrt{2\pi S(I)}} \exp\left(\frac{-\frac{1}{2}[L(I)\phi - A(I)]^2}{S(I)}\right)$$

where $L(I)$, $A(I)$, and $S(I)$ are as described above, we obtain compatible ϕ-measures, provided we can find compatible projections satisfying (1.2).

If $L(I)$ is an absolutely continuous additive set function relative to a measure $\mu(s)$, it is well known that it has the form

$$L(I) = \int_I \psi(s)d\mu(s) \ .$$

If we then take

$$\phi_\nu = \int_{I_\nu} \frac{\phi(s)\psi(s)d\mu(s)}{L(I_\nu)}$$

and define $P\phi$ to be the piecewise constant function which is ϕ_ν on I_ν, (1.2) is satisfied and the projections are compatible.

The special case with

$$A(I) = 0, \ L(I) = \mu(I), \ S(I) = \frac{1}{2a}\,\mu(I)$$

is precisely the one which was presented in Chapter I.

Example II: The Heaviside distributions

The structurally simplest family which is closed under convolution is that generated by the Heaviside function

$$\eta(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

which has mean 0 and variance 0. It is easily verified that $\eta(x-c)$ has mean c and variance 0, and that

$$\eta(x-c_1) * \eta(x-c_2) = \eta(x - (c_1+c_2)).$$

Hence, in this case we are forced to take $S(I) = 0$. If we assume $A(I)$ is absolutely continuous relative to a measure $\mu(s)$ and set

$$A(I) = \int_I \psi(s)d\mu(s)$$

and take

$$L(I) = \int_I d\mu(s)$$

(where $\psi(s)$ is a fixed function which is $d\mu(s)$-integrable), the projections into step functions (as in example II) are again compatible. In this case the resulting integral may be written as

$$\int^{\infty} f[\phi] \prod_{s} d\eta(\phi(s)-\psi(s))$$

and under a mild hypothesis on f can be proved to be equal to $f[\psi]$.

Example III: Classical Poisson Distribution

The classical Poisson distribution is given by

$$m(x,c) = \sum_{n \leq x} \frac{c^n}{n!} e^{-c} \qquad (c > 0)$$

and has mean and variance equal to c. It can be verified that

$$m(x,c_1) * m(x,c_2) = m(x,c_1+c_2) .$$

Again the choice of the set functions A(I) and S(I) is restricted. Here we must have $a(I) = \sigma^2(I)$, which implies

(4.5) $$L(I) = S(I)/A(I)$$

and hence

$$a(I) = \sigma^2(I) = \frac{A^2(I)}{S(I)}$$

Our ϕ-measure will then be of the form

$$m_I(x) = \sum_{\phi \leq x} \frac{1}{\phi!} \left[\frac{A^2(I)}{S(I)}\right]^{\phi} \exp\left[\frac{-A^2(I)}{S(I)}\right]$$

(where the summation is taken over all integral values of ϕ).

Since L(I) must satisfy (4.5) there is, in general, some difficulty in the construction of the corresponding compatible

projections and, indeed, this may not always be possible. The case discussed previously in the lectures corresponds to the choice $S(I) = A(I) = \mu(I)$, and $L(I) = 1$.

The following examples may also be worked out along similar lines.

Example IV: Generalized Poisson Distribution

This is given by

$$m(x,a,b,c) = \sum_{an+b \leq x} \frac{c^n}{n!} e^{-c} \qquad (c > 0) .$$

For fixed a, as a family of distributions depending on the parameters b, c, these are closed under convolution. In fact for $c_1, c_2 > 0$

$$m(x,a,b_1,c_1) * m(x,a,b_2,c_2) = m(x,a,b_1+b_2,c_1+c_2) .$$

The mean and variance of this distribution are $(b+c)/a$ and c/a^2 respectively.

Example V: Binomial Distribution

For $n = 1,2,\ldots$ and $0 < p < 1$ let

$$m(x,a,n,p) = \sum_{\nu \leq x-c} \binom{n}{\nu} p^{\nu} (1-p)^{n-\nu} .$$

Then we have

$$m(x,a_1,n_1,p) * m(x,a_2,n_2p) = m(x,a_1+a_2,n_1+n_2,p) .$$

The mean of this distribution is $(a+np)$ and the variance is $[(a+np)^2 + np(1-p)]$.

Example VI: The Pearson Distribution

For

$$m(x,a,\lambda) = \begin{cases} 0 & x < 0 \\ \dfrac{a^{\lambda}}{\Gamma(\lambda)} \displaystyle\int_0^x t^{\lambda-1} e^{-at}\, dt & x \geq 0 \end{cases}$$

we have

$$m(x,a,\lambda_1) * m(x,a,\lambda_2) = m(x,a,\lambda_1+\lambda_2) .$$

Chapter IV

Integration Spaces

§ 1. Introduction.

As the next step in the development of our discussion of integration of functionals, we wish to set down an abstract structural setting within which integrals such as have been discussed (and will be considered) may be formulated. The composite object which we consider will be called an integration space.

We assume the following entities to be given:

(1) a linear space G of elements x;

(2) a collection $\mathfrak{R} = \{R\}$ of linear subspaces R of G and a partial ordering $<$ of the subspaces $R \in \mathfrak{R}$;

(3) for each $R \in \mathfrak{R}$ a linear projection P_R of G onto R, (projecting each element x of G into $P_R x$); and

(4) corresponding to each $R \in \mathfrak{R}$ a Borel field $\mathfrak{B}(R)$ of subsets of R and a totally additive measure m_R over $\mathfrak{B}(R)$.

We further assume that the above entities have the following properties:

(α) for $R_1, R_2 \in \mathfrak{R}$ there is an $R \in \{\mathfrak{R}\}$ such that

$$R > R_1 \ , \quad R > R_2$$

(β) if $R' > R$ and $R, R' \in \{\mathfrak{R}\}$, then

$$P_R P_{R'} x = P_R x$$

for all $x \in G$.

(γ) the measures on the subspaces R of $\mathfrak{R}$ are compatible; that is, suppose $R', R \in \mathfrak{R}$ with $R' < R$, and $S \in \mathfrak{B}(R)$. Let $S' = \{x \in R' : P_R x \in S\}$. Then we have

(i) $$S' \in \mathfrak{B}(R')$$

(ii) $$m_{R'}(S') = m_R(S) \ .$$

§ 2. Cylinder Sets.

We define a cylinder set, or simply a cylinder, in G to be a subset C of G of the form

$$C = \{x \in G: \ P_R x \in S\}$$

for some $R \in \mathcal{R}$ with $S \in \mathcal{S}(R)$. The set S is then referred to as a base of the cylinder C, or more explicitly as the base in R of this cylinder.

Lemma: A cylinder C with a base in $R \in \mathcal{R}$ is also a cylinder with respect to a base in any $R' \in \mathcal{R}$ such that $R' > R$. In fact, let S be the base in R. Then $S' = \{x \in R': \ P_R x \in S\}$ is the desired base in R'.

Proof: The assumption (γ) in § 1 provides that $S' \in \mathcal{S}(R')$, so that it suffices to prove that

$$x \in C \Leftrightarrow P_{R'} x \in S' .$$

This follows from

$$x \in C \Leftrightarrow P_R x \in S \Leftrightarrow P_R P_{R'} x \in S \Leftrightarrow P_{R'} x \in S' \quad .$$

Using the above lemma, one easily establishes the following facts:

(a) the null set and the whole space G are cylinders;
(b) the intersection of a finite number of cylinders is a cylinder;
(c) the union of a finite number of cylinders is a cylinder; and
(d) the complement of a cylinder is a cylinder.

To a cylinder C with base S in $R \in \mathcal{R}$ we assign the measure

$$m(C) = m_R(S) . \tag{2.1}$$

It is a consequence of the compatibility condition (γ) that m(C) is defined over cylinders, independently of the choice of base S. This together with the facts listed above yields that m(C) is a finitely additive measure over the set of cylinders.

§ 3. Cylinder functions.

A function $f[\cdot]$ is called a cylinder function if there exists an $R \in \mathcal{R}$ such that

$$f[P_R x] = f[x]$$

The cylinder function is then said to have a base in R.

If f has a base in $R \in \mathcal{R}$, and $R' \supset R$, $R' \in \mathcal{R}$, then f has a base in R'. For, by (β),

$$f[P_{R'} x] = f[P_R P_{R'} x] = f[P_R x] = f[x] .$$

From this remark and property (α) it is easily inferred that the collection of cylinder functions forms a linear space.

Lemma: If f is a cylinder function with base in $R \in \mathcal{R}$, and is measurable as a function over R then, for any $R' \in \mathcal{R}$ such that $R' > R$, f is measurable over R' and

$$\int_{R'} f[x] dm_{R'} = \int_R f[x] dm_R$$

Proof: First we prove the measurability over R'. Let

$$S = \{x \in R: \ f[x] < a\}$$

so that $S \in \mathcal{A}(R)$. Set $S' = \{x \in R': \ f[x] < a\}$. We wish to show that $S' \in \mathcal{A}(R')$. This follows from the fact that

$$x \in S' \Leftrightarrow f[x] = f[P_R x] < a, \ x \in R'$$
$$\Leftrightarrow P_R x \in S, \ x \in R'$$

so that S' is precisely the same as in assumption (γ) of §1, and the desired result follows.

As for the equality of the integrals, we need only consider the case of $f[x] \geq 0$. Then we form

$$\sigma = \sum_i y_i \, m_R \{x \in R: \ y_i \leq f[x] < y_{i+1}\} \quad \text{and}$$
$$\sigma' = \sum_i y_i \, m_{R'} \{x \in R': \ y_i \leq f[x] < y_{i+1}\} .$$

As in the preceding paragraph we obtain

$$\left\{x\varepsilon R': y_i \le f[x] < y_{i+1}\right\} = \left\{x\varepsilon R': P_R x\varepsilon\left\{x\varepsilon R: y_i \le f[x] < y_{i+1}\right\}\right\}$$

so that $\sigma = \sigma'$ and the desired result follows.

Thus, if a cylinder function having a base in $R\varepsilon\mathcal{R}$ is integrable over R, it has the same integral over all subspaces in $\mathcal{R}$ with respect to which it has a base. (See the Lemma earlier in this section.) For such a cylinder function f we may define the integral

$$I^{\mathcal{R}}(f) = \int_R f[x]dm_R$$

which is independent of the choice of $R\varepsilon\mathcal{R}$ as base space. The dependence of the integral $I^{\mathcal{R}}(\cdot)$ on the collection $\mathcal{R}$ of subspaces R is indicated by a superscript.

§ 4. The spaces $\mathcal{L}_a^{\mathcal{R}}$

For any fixed $a \ge 1$ we consider those cylinder functions f for which $|f|^a$ is integrable with respect to a base $R\varepsilon\mathcal{R}$. These, clearly, form a linear space $\mathcal{C}_a^{\mathcal{R}}$ and we may introduce the norm

$$\|f\|_{\mathcal{R}} = \|f[\cdot]\|_{\mathcal{R}} = \left[I^{\mathcal{R}} (|f|^a)\right]^{\frac{1}{a}} .$$

Furthermore, we may complete $\mathcal{C}_a^{\mathcal{R}}$ by introducing ideal functions $\hat{f}[\cdot]$ which are associated with a sequence of cylinder functions $f_n[\cdot]\varepsilon\mathcal{C}_a^{\mathcal{R}}$ such that

$$\|(f_n - f_m)\|_{\mathcal{R}} \to 0 \qquad \text{as } m,n \to \infty .$$

Since

$$\left| \|f_n\|_{\mathcal{R}} - \|f_m\|_{\mathcal{R}} \right| \le \|f_n - f_m\|_{\mathcal{R}} \to 0 ,$$

$\lim_{n\to\infty} \|f_n\|_{\mathcal{R}}^*$ exists for such sequences $\{f_n[\cdot]\}$ and we may define

$$\|\hat{f}\|_{\mathcal{R}} = \lim_{n\to\infty} \|f_n\|_{\mathcal{R}}^a$$

and

$$I^{\mathcal{R}}(|\hat{f}|^a) = \left[\|\hat{f}\|_{\mathcal{R}} \right]^a .$$

The space of these ideal functions will be denoted by $\mathcal{L}_a^{\mathcal{K}}$.

We note that the space $\mathcal{L}_2^{\mathcal{K}}$ is a Hilbert space with inner product given by

$$(\hat{f},\hat{g}) = I^{\mathcal{K}}(\hat{f}\cdot\hat{g})$$

§ 5. Examples of Integration spaces.

Special realizations of the notion of integration space set forth in the preceding sections are quite numerous. We consider first the example in which G consists of the unrestricted infinite sequences of real numbers

$$x = (x_1, x_2, \ldots) .$$

For $\mathcal{K}$ we take the finite dimensional subspaces

$$R_m: \quad (x_1,\ldots,x_n,0,0,\ldots) \tag{5.1}$$

with projections P_n given by

$$P_n x = (x_1,\ldots,x_n,0,0,\ldots) \tag{5.2}$$

As measure in R_n we may, for example, take the Gaussian measure given by

$$dm_{R_n} = (\sqrt{2\pi})^{-n} \exp\left[-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right]dx_1\ldots dx_n \tag{5.3}$$

All of the requirements set down in §1 are then easily verified and the formalism which has been developed applies here.

Another example which fits into this scheme is the procedure for integrating functionals given in Chapter I. In that procedure G consists of all functions $\phi(s)$ defined for s in a given base space $\mathcal{S}$. The collection consists of the subspaces R_ρ of those functions which are piecewise constant over the intervals of a partition ρ of the interval $\mathcal{S}$. The projections P_R may be defined by

$$P_R\phi = \left\{\phi_\nu = \frac{\int_{\mathcal{S}_\nu}\phi(s)ds}{\int_{\mathcal{S}_\nu} ds} \quad \text{on } \mathcal{S}_\nu \qquad (\nu = 1,\ldots,n)\right\}$$

where the $\mathcal{I}_\nu$ are the intervals of the partition ρ. As measure over R_ρ we take the Gaussian measure given by

$$dm_{R_\rho} = (\sqrt{2\pi})^{-n} \exp\left[-\frac{1}{2}\sum_{\nu=1}^{n} \phi_\nu^2\right] d\phi_1,\ldots d\phi_n$$

the order relation > is inclusion and $\mathcal{M}(R)$ consists of the Gauss-measurable subsets of R. That the conditions set down in §1 are then satisfied has already been verified in previous lectures.

§ 6. Total additivity of measure.

The question arises of whether or not the measure defined above over cylinders is totally additive. Indeed, in the case of the example of the space of unrestricted sequences this is so; as follows from a theorem of Kolmogorov. (A. Kolmogorov, Foundations of the Theory of Probability, p. 17). In fact, in this case, the measure over cylinders may be extended to a totally additive measure over the smallest Borel field which contains the cylinders.

That the measure over cylinders need not in general be totally additive is easily seen. One need only choose for G a subspace of the space of unrestricted sequences which is Borel measurable in that theory and such that it has measure 0. Also, if we arrange that G contains the R_n of (5.1), we can use these for $\mathcal{R}$ and together with the projections of (5.2) and measures (5.3) develop the "G-theory". Since G has measure 0 in the unrestricted sequence theory we can cover G with a countable number of disjoint cylinders whose total measure is as small as we like. But the measure of a cylinder will be the same in the G-theory and the unrestricted sequence theory. Since on the other hand the measure of G must be 1 in the G-theory we see that the measure over cylinders is not totally additive in the G-theory.

We shall return to these ideas in later sections. In particular when we treat the case in which G is a separable Hilbert space, an explicit example will be given which demonstrates the lack of total additivity.

Chapter V
Integration over Hilbert Space

§ 1. Introduction.

We now propose to develop in detail a notion of integration over a separable real Hilbert space, within the framework of the general concept of "integration space". Taking advantage of the special structure of such a Hilbert space, together with the "unitary invariance" of the Gaussian measure, we are led to the derivation of a certain kind of "unitary invariance" of this integral. In addition, the spaces $\mathcal{L}_a^{\mathfrak{k}}$ will have certain interesting special features. In particular, some statement can be made concerning the association of "proper functions" with "ideal functions".

§ 2. Notions and Notations.

The following notations will be used:

(a) $\mathfrak{k}$ a separable real Hilbert space, with points x, inner product (x,y), norm $\|x\|$.

(b) P,Q,R,S,T,etc. denote orthogonal projections (henceforth simply called projections) into finite dimensional subspaces of $\mathfrak{k}$. We will also use a symbol such as P to denote the range of the projection P.

(c) $(P \oplus Q)$ denotes the linear sum of the ranges of P and Q; by PQ we denote - as is customary - the operator consisting in applying first Q and then P.

(d) The formula $R \sim (P \oplus Q)$ is to indicate that the operator R is the projection onto the space $(P \oplus Q)$; hence, by the convention set down in (b), R may also be used as another name for this space. The formula $S \sim PQ$ is to indicate that S is the projection onto the range of the operator PQ.

(e) dim M is the dimension of the range of the operator M.

In the case of Hilbert space, a finite dimensional subspace uniquely determines its orthogonal projection, and conversely, so that little ambiguity can arise from the use of P as a notation both for a projection, and for the range of this projection.

Definitions:

V1. A sequence of orthogonal projections $\mathcal{K} = (P^n)$ $(n=1,2,\ldots)$ will be called monotone if, for $n > m$, $P^n \supset P^m$.

V2. A sequence of orthogonal projections (P^n) will be called complete if, for all $x \in \mathcal{h}$, $P^n x \to x$ as $n \to \infty$; that is, $\|P^n x - x\| \to 0$.

V3. A set $\mathcal{K} = (R)$ of orthogonal projections will be called complete if it contains at least one complete subsequence.

V4. A set $\mathcal{K} = (R)$ of orthogonal projections will be called closed if for every given pair of projections $P, Q \in \mathcal{K}$ there exists an $R \in \mathcal{K}$ such that $P \subset R$ and $Q \subset R$.

V5. A sequence (P^n) of orthogonal projections will be called a complete basic system if it is complete, monotone, and, for each n, $\dim P^n = n$.

Thus to a complete basic system (Q^n) there corresponds an orthonormal basis $\{g_i\}$ of $\mathcal{h}$ such that, for $x = \sum_{i=1}^{\infty} x_i g_i$,

$$Q^n x = \sum_{i=1}^{n} x_i g_i \quad , \quad n = 1,2,\ldots$$

Clearly, any infinite subsequence of a complete basic system is complete and monotone. Furthermore, by the insertion of certain projections a given complete monotone sequence may be converted into a complete basic one.

Frequently, we shall omit the superscript and write (Q) in place of (Q^n).

A particular complete system which will be of interest is the collection $\hat{\mathcal{K}}$ consisting of all projections into finite dimensional subspaces of $\mathcal{h}$.

§ 3. $\mathcal{h}$ as an integration space.

In treating $\mathcal{h}$ as an integration space we take as subspaces those corresponding to a given complete system $\mathcal{K}$ of projections. Note that condition (α) of IV §1 is satisfied by virtue of V1. while condition (β) is satisfied since the projections P are assumed orthogonal.

The measures will be taken as the usual Gaussian measures, i.e. for any $P \in \mathfrak{L}$

$$(3.1) \qquad dm_P = \sqrt{2\pi}^{-K} \exp\left[-\tfrac{1}{2}\|Px\|^2\right] dx_1 \ldots dx_K$$

where K = dim P. These of course correspond to the Gaussian distribution of mean 0 and variance 1. The analogous measures which correspond to the Gaussian distribution of mean 0 and variance σ^2 will be discussed later. In view of the obvious unitary invariance of this measure, it is determined independently of the particular choice of $\{x_i\}$ as a basis for P. Instead of (3.1) we shall often write

$$(3.2) \qquad dm_P = \exp\left[-\tfrac{1}{2}\|Px\|^2\right] d_P x$$

It is then clear that also property (γ), required of an integration space, is satisfied. Because of a special feature of the case we are considering here we shall elaborate somewhat on the general construction.

Denote by $\mathfrak{Z}_a^P$ $(a \geq 1)$ the linear space of those cylinder functions f(x) with base in P such that $I^P(|f|^a)$ exists. I.e., such that $|f|^a$ is integrable over P with respect to the measure dm_P, or simply P-integrable. Note: if f(x) is a cylinder function with base in P which is P-measurable, then f(Qx) is a cylinder function with base in Q which is Q-measurable. To verify this last assertion, form $R \sim (P \oplus Q)$ and recall that the set

$$S = \{x \in R: \ f(x) < \beta\}$$

is R-measurable. Then, since Q itself is an R-measurable subset of R, the set $(Q \cap S)$ is R-measurable from the unitarily invariant nature of the measures involved here. It then follows that

$$Q \cap S = \{x \in Q: \ f(x) < \beta\}$$

is Q-measurable, and the desired result ensues.

In constructing the <u>ideal functions</u> $\hat{f}(x) = \{f_n(x)\}$ of $\mathfrak{Z}_a^P$

for a fixed $a \geq 1$, we shall admit sequences $\{f_n(x)\}$ such that $f_n(x)$ is a cylinder function with base in $P^n \mathfrak{k}$ and

(a) $f_n(x) \in \mathcal{L}_a^{P^n}$ for n sufficiently large,

(b) $\lim_{m,n\to\infty} \|f_m - f_n\|_{\mathfrak{k}} = 0$.

The remainder of the setup is then exactly as described in the general discussion of integration spaces.

We shall find it convenient to make use of the following notion:

Definition: Let P be a projection, and $f(\cdot)$ a function over $\mathfrak{k}$; then we define the function $Pf(\cdot)$ by

$$Pf(x) = f(Px) \quad .$$

For a given complete monotone system (P) the function f(x) is (P)-integrable, or $f(x) \in \mathcal{L}_a^{(P)}$ if the sequence $\{P^n f\}$ is an ideal element of $\mathcal{L}_a^{(P)}$. We then define

$$\|f\|_{(P)} = \lim_{n\to\infty} \|P^n f\|_{(P)} \quad .$$

The collection of those "proper functions" which are (P)-integrable clearly forms a linear space which we also denote by $\mathcal{L}_a^{(P)}$

In §5, we shall extend the above, and give a definition of $\mathfrak{K}$-integrability for any complete closed set $\mathfrak{K}$ of projections.

§ 4. The non-total additivity of the measure over cylinders

As was already mentioned in §6 of the general discussion of integration spaces in Chapter IV, the measure over cylinders, determined by the above integration theory over a Hilbert space, is not totally additive. In order to deduce this from our previous remarks on this subject, it suffices to consider the sequential Hilbert space $\mathfrak{k}$ of sequences

$$x = (x_1, x_2, \ldots x_i, \ldots), \quad \sum_{i=1}^{\infty} x_i^2 < \infty$$

as a subset of the space $\sum$ of unrestricted sequences. Within the integration theory of $\sum$ (with Gaussian measures), the measure over cylinders is totally additive and may be extended to a totally additive measure on the smallest Borel field containing the cylinders. It is easily seen that $\mathcal{K}$ is a Borel measurable set in the $\sum$ theory. All that need be shown then is that the measure of $\mathcal{K}$ is 0 in the $\sum$ theory. From the "Zero or One Law"* of probability theory, this measure is 0 or 1, and that is indeed 0 is an immediate consequence of the "Three Series Theorem" of Khinchine and Kolmogoroff; (in fact, the Two Series Lemma* is sufficient). To see this fact directly, consider the subset J of $\sum$ consisting of those x such that $x_n < 1$ for infinitely many values of n. Clearly, $\mathcal{K} \subset J$ and the measure of J in the $\sum$ theory is 0 so the measure of $\mathcal{K}$ in this theory is also 0.

This suggests the possibility of giving a simple example in which $\mathcal{K}$ is covered by the union of a countable number of cylinders, $\mathcal{K} \subseteq \bigcup_{i=1}^{\infty} C_i$, such that $\sum_{i=1}^{\infty} m(C_i) < 1$. This may be done as follows: choose δ such that

$$\frac{1}{\sqrt{2\pi}} \int_{-\delta}^{\delta} \exp\left[-\frac{1}{2} u^2\right] du = \varepsilon < \frac{1}{2}$$

and form the cylinders

$$C_1 = \left\{x: \ |x_1|, \ |x_2| < \delta\right\}$$
$$C_2 = \left\{x: \ |x_3|, \ |x_4| < \delta\right\}$$

and, in general,

$$C_n = \left\{x: \ |x_{(2^{n-1}+1)}|, \ldots, |x_{2^n}| < \delta\right\}$$

Clearly we then have that $\mathcal{K} \subseteq \bigcup_{n=1}^{\infty} C_n$. Since

---

* Kolmogoroff, Foundations of the Theory of Probability.

$$m(C_n) < \epsilon^{(2^{n-1})} \leq \epsilon^n \quad ,$$

we have

$$\sum_{n=1}^{\infty} m(C_n) < \frac{\epsilon}{1-\epsilon} < 1 \ .$$

In fact, we see from the above that we can have an example in which $\sum_{n=1}^{\infty} m(C_n)$ is as small as we like.

Though the C_i in the above example are not mutually disjoint, an example with mutually disjoint cylinders may be obtained by "disjointing" the C_i in the standard way.

§ 5. Invariance and Identification.

The general problem of "identification" is two-fold in nature. On one hand we have the problem of ascertaining some "reasonable" sense in which for any two complete basic systems (P) and (Q) an ideal function of $\mathcal{L}_a^{(P)}$ may be "identified" with an ideal function of $\mathcal{L}_a^{(Q)}$. On the other hand, we have the closely related problem of prescribing how to associate with a "proper" function f(x) an ideal function of an $\mathcal{L}_a^{(P)}$. Several notions, related to these questions will now be proposed and discussed.

Definitions: Invariance Properties

(I) Semi-invariance relative to $\mathcal{L}$

A proper function f(x) is called semi-invariant relative to a given complete closed system $\mathcal{L}$ if, for every complete monotone subsequence (Q) = (Q^n) of $\mathcal{L}$,

(a) $f(x) \in \mathcal{L}_a^{(Q)}$

(b) $\| Qf \|_{(Q)}$ is independent of (Q) .

For a semi-invariant function, since (b) is satisfied we shall write $\|f\|_{\mathcal{L}} = \|Qf\|_{(Q)}$.

(II) Invariance relative to $\mathcal{L}$

A proper function f(x) is called invariant relative to a given complete closed system $\mathcal{L}$ if, for every pair of

complete monotone subsequence of $\mathcal{L}$, $(P) = (P^n)$ and $(Q) = (Q^m)$, we have

$$P^n f(x) = f(P^n x) \in \mathcal{L}_a^{P^n} \quad \text{and} \quad Q^m f(x) = f(Q^m x) \in \mathcal{L}_a^{Q^m},$$

for n and m sufficiently large and

$$\lim_{n,m\to\infty} \| P^n f - Q^m f \|_{(P^n \dot{+} Q^m)} = 0 \tag{5.1}$$

(III) Complete invariance relative to $\mathcal{K}$

A proper function $f(x)$ is called completely invariant relative to a given complete closed system $\mathcal{K}$ if, for every complete subsequence of projections (Q^n) of $\mathcal{K}$ we have $Q^n f \in \mathcal{L}_a^{Q^n}$ for sufficiently large n, and

$$\lim_{n,m\to\infty} \| Q^n f - Q^m f \|_{\mathcal{K}} = 0$$

We may now introduce the notion of $\mathcal{R}$-integrability, as provised at the conclusion of §3.

Definition: Given a complete closed system $\mathcal{R}$, a function f will be called $\mathcal{R}$-integrable, or, more generally, $f(x) \in \mathcal{L}_a^{\mathcal{R}}$ $(a \geq 1)$ if $f(x)$ is invariant relative to $\mathcal{R}$.

If $f(x)$ is $\mathcal{R}$-integrable it is easily seen that $\|f\|_{(P)}$ is the same for any choice of a complete monotone subsequence (P) of $\mathcal{R}$. We then define $\|f\|_{\mathcal{R}} = \|f\|_{(P)}$, and introduce the $\mathcal{R}$-integral of $|f|^a$ over $\mathcal{K}$ as

$$\mathcal{R}^I_{\mathcal{K}} (|f|^a) = [\|f\|_{\mathcal{R}}]^a .$$

We note that the complete closed system $\mathcal{R}$ has the property that any complete monotone subsequence may be filled out to a complete basic subsequence inside $\mathcal{R}$. Then, in verifying the first two of the above notions of invariance, it suffices to consider only complete basic subsequences of $\mathcal{R}$.

We now gather together a few simple observations several of which will play an important role in later discussions inasmuch as they will be used in proving our main theorem.

Remarks:

(5.2) The notions of invariance as listed above are in order of increasing strength in that (III) implies (II) implies (I).

(5.3) The invariant and completely invariant proper functions form linear spaces.

(5.4) If f(x) and g(x) are both invariant (completely invariant) proper functions, and if both are bounded, then f(x)g(x) is invariant (completely invariant).

(5.5) Suppose that f(x) is a cylinder function with base in P belonging to $\mathcal{L}_a^P$. Then the complete invariance of f is equivalent to the assertion that, for every sequence (Q^n) such as is described in (III), firstly $Q^n f \in \mathcal{L}_a^{Q^n}$ for n sufficiently large and secondly, for $R^n \sim (P \oplus Q^n)$,

$$\lim_{n\to\infty} \|f - Q^n f\|_{R^n} = 0 .$$

(5.6) A natural question which arises - and which is not, as yet, settled - is: whether or not any of the apparently weaker notions of invariance will necessarily imply a stronger one. In particular one might ask whether or not semi-invariance implies invariance. A partial result in this direction will be obtained. However, since it stems from a theorem which will be proved in a later section we shall postpone further consideration of this question to §8.

(5.7) The main result concerning invariance is:

Theorem A. Every cylinder function is completely invariant.

It will be proved in § 7. At present we continue to introduce definitions.

Definitions: Identification

(I) Semi-identification

Let (P) be a complete basic system, and consider any $\hat{f}(x) = \{f_n(x)\} \in \mathcal{L}_a^{(P)}$. Then, as has already been noted, $\|\hat{f}\|_{(P)}$ is defined and is in fact given by

$$\|\hat{f}\|_{(P)} = \lim_{n\to\infty} \|f_n\|_{(P)} .$$

We wish then to "identify" $\hat{f}(x)$, in a natural way, with an ideal element of <u>each</u> $\mathfrak{L}_a^{(Q)}$, where (Q) is a complete basis system, in such a way that $\|\hat{f}\|_{(P)}$ may be evaluated in the (Q)-theory. This idea may be formalized by the notion of <u>semi-identification</u>. This is said to be possible if, for every two complete basic systems (P) and (Q), given any $\hat{f}(x) = \{f_n(x)\} \varepsilon \mathfrak{L}_a^{(P)}$ there exists a sequence of integers m_n (n=1,2,3,...) such that

$$\text{(a)} \qquad \{Q^{m_n} f_n\} \varepsilon \mathfrak{L}_a^{(Q)}$$

$$\text{(b)} \qquad \|\{Q^{m_n} f_n\}\|_{(Q)} = \|\{f_n\}\|_{(P)} \ .$$

(II) <u>Identification</u>

Though semi-identification provides that $\|\hat{f}\|_{(P)}$ may be evaluated with reference to any complete basic system (Q), it does not imply that, for any complete basis (Q), the space $\mathfrak{L}_a^{(Q)}$ is a "faithful representation" of $\mathfrak{L}_a^{R}$. If this is the case we shall speak of (unqualified) <u>identification</u>. More precisely, identification is said to be possible if, for every two complete basic systems (P) and (Q), given any $f_n(x) \varepsilon \mathfrak{L}_a^{(P)}$ there exists a sequence of integers m_n (n=1,2,3,...) such that

$$\lim_{n\to\infty} \|f_n - Q^{m_n} f_n\|_{\hat{\mathfrak{K}}} = 0 \ .$$

<u>Remarks</u>.

(5.8) If identification is possible, so is semi-identification.

(5.9) As a simple consequence of Theorem A, one may derive

<u>Theorem B.</u> Identification is possible.

<u>Proof</u>: From Theorem A it follows that cylinder functions are invariant. Hence, if $\{f_n(x)\} \varepsilon \mathfrak{L}_a^{(P)}$ where the f_n are cylinder functions; and (Q) is any complete basic system, we have

$$\lim_{m\to\infty} \|f_n - Q^{m} f_n\|_{\hat{\mathfrak{K}}} = 0$$

for each n. Therefore we may choose m_n such that

$$\|f_n - Q^{m_n} f_n\|_{\hat{\mathcal{R}}} < 2^{-n}$$

and hence

$$\lim_{n\to\infty} \|f_n - Q^{m_n} f_n\|_{\hat{\mathcal{R}}} = 0 .$$

§ 6. The Fundamental Inequality

The heart of the proofs of Theorem A and related statements and, indeed, the essence of the whole theory is contained in the following

Main Lemma: Suppose f(x) is a cylinder function with base in P, $f(x) \in \mathcal{L}_a^P; \{Q^n\}$ is any simply complete sequence of projections. Then, for n sufficiently large,

$$\|Q^n f\|_{Q^n} \leq 2 \|f\|_P \quad . \tag{6.1}$$

Proof: The first step is to show that $\|Qf\|_Q$ may always be calculated in a space of dimension $\leq$ dim P. Set $S \sim QP$ (see V§1) so that in particular $S \subset Q$, dim $S \leq$ dim P, and SQP = QP.

Since

$$(x, PQSy) = (SQPx, y) = (QPx, y) = (x, PQy)$$

we have

$$PQS = PQ \tag{6.2}$$

Since $S \subset Q$, we have QS = S, so that (6.2) implies

$$PS = PQS . \tag{6.3}$$

Hence, since f(x) = f(Px), we have

$$f(Qx) = f(PQx) = f(PSx) = f(Sx) ,$$

so that f(Qx) = f(Sx) is seen to be a cylinder function with base in S. Consequently,

$$\|Qf\|_Q = \|Sf\|_S \quad .$$

Next we have

$$(6.4) \qquad \|Sf\|_S = \int_S |f(PSx)|^a \exp[-\tfrac{1}{2}\|x\|^2] d_S x$$

$$\leq \int_S |f(Px)|^a \exp[-\tfrac{1}{2}\|Px\|^2] d_S x \quad .$$

Since $PS = PQ$ we have $SP = QP$. Therefore, since $Q^n x \to x$ as $n \to \infty$, we have $\dim Q^n P \to \dim P$ and, hence, for $S^n \sim Q^n P$,

$$\dim S^n P \to \dim P \quad .$$

Consequently, for n sufficiently large, we have $\dim S^n P = \dim Q^n P = \dim P$. Since $\dim S^n \geq \dim S^n P$ we have, therefore, $\dim S^n \geq \dim P$ for n sufficiently large. On the other hand, $\dim S^n \leq \dim P$. Hence we have

$$\dim S^n = \dim P \quad .$$

Assuming n is large enough we are in a position to make a change of variable in (6.4); $y = Px$ for $x \in S^n$.

Fixing a basis for P independently of n and choosing bases for each S^n, we may represent the transformation $y = Px$, $x \in S^n$, by a matrix. Let Δ_n be the determinant of this matrix. Then the integral in (6.4) becomes

$$|\Delta_n^{-1}| \int_P |f(y)|^a \exp[-\tfrac{1}{2}\|y\|^2] d_P y \quad .$$

Now, since Δ_n is the determinant of the matrix representing the transformation PS^n of S^n into P, it is also the determinant of the transposed transformation $S^n P$ of P into S^n. Since

$$(PS^n)(S^n P) = PSP = PQ^n P \quad ,$$

we have

$$\Delta_n^2 = \det (PS^n) \det (S^n P) = \det (PQ^n P) \quad .$$

Now, as $n \to \infty$, $PQ^nP \to P$ so that $\Delta_n^2 \to 1$. Hence $|\Delta_n^{-1}| \to 1$ and (6.1) follows for n sufficiently large.

Corollary: If f(x) is a cylinder function with base in P such that $f(x) \in \mathcal{L}_a^P$ then, for all sufficiently large m,

$$f(Q^n x) \in \mathcal{L}_a^n \quad .$$

§ 7. Proof of Theorem A.

In carrying out the proof of Theorem A it is most convenient to make use of the following notion.

Definition: A set $\mathcal{C}$ of cylinder functions in $\mathcal{L}_a^R$ is said to be dense if, for any given cylinder function with base in P such that $f(x) \in \mathcal{L}_a^P$, and, for any given $\varepsilon > 0$, there is a cylinder function g(x) with base in P, $g(x) \in \mathcal{C} \cap \mathcal{L}_a^P$, such that

$$\|f - g\|_P < \varepsilon \quad . \tag{7.1}$$

Lemma 7.1. If there exists a dense set $\mathcal{C}$ of cylinder functions each of which is completely invariant then Theorem A holds.

Proof: Let f(x) be a cylinder function with base in P, $f(x) \in \mathcal{L}_a^P$, and let (Q^n) be any simply complete sequence of projections. Set $R^n \sim (P \oplus Q^n)$. Given $\varepsilon > 0$, choose a function g(x) in $\mathcal{C}$ with base in P, $g(x) \in \mathcal{L}_a^P$, such that

$$\|f - g\|_P < \varepsilon \quad . \tag{7.2}$$

Then, by the above corollary, for n sufficiently large, $f(Q^n x)$, $g(Q^n x) \in \mathcal{L}_a^{Q^n}$. Hence by the fundamental inequality (6.1) we get

$$\|Q^n f - Q^n g\|_{Q^n} < 2\varepsilon \quad . \tag{7.3}$$

Also, for n sufficiently large, since g(x) is completely invariant,

$$\|g - Q^n g\| < \varepsilon \quad . \tag{7.4}$$

Then combining (7.2), (7.3), and (7.4) we have

$$\|f - Q^n f\|_{R^n} \leq \|f - g\|_{R^n} + \|g - Q^n g\|_{R^n} + \|Q^n g - Q^n f\|_{R^n}$$

$$= \|f - g\|_{P^n} + \|g - Q^n g\|_{R^n} + \|Q^n g - Q^n f\|_{Q^n}$$

$$< 4\varepsilon \quad .$$

Hence, it follows from remark (5.5) that f(x) is completely invariant.

Thus to complete the proof of Theorem A it suffices to prove

Lemma 7.2. There exists a dense set of cylinder functions each of which is completely invariant.

Proof: We maintain that it suffices to provide for each P with dim P = 1 a set of bounded cylinder functions with base in P, each in $\mathcal{L}_a^P$, and dense in $\mathcal{L}_a^P$, such that each is completely invariant.

Consider the bounded functions $\chi[t]$ satisfying a Lipschitz condition,

$$|\chi[t] - \chi[t']| \leq c|t - t'| \quad .$$

Let g denote elements of P, with $\|g\| = 1$. The set of functions $\{(g,x)\}$ is dense in $\mathcal{L}_a^P$ for, according to remarks (5.3) and (5.4), sums and products of such functions are completely invariant and the functions generated by such products are dense, as is well known. We next show that each of these is completely invariant. According to remark (5.5), it is sufficient to prove the relation

$$\|\chi[(g,x)] - \chi[(g,Q^n x)]\|_{R^n} \to 0 \quad \text{as } n \to \infty$$

where (Q^n) is a sequence as described in III and $R^n \sim (P \oplus Q^n)$.

Instead of R^n we first introduce the subspaces

$$T^n \sim (P \oplus Q^n P)$$

whose dimension is bounded: dim $T^n \leq 2$ dim P. Clearly, we have

$$PQ^nT^n = PQ^n$$

and

$$PT^n = P \quad .$$

Hence, for g and x in P, we have

$$(g,x) = (g,PT^nx) = (g,T^nx) \quad ,$$
$$(g,Q^nx) = (g,PQ^nx) = (g,PQ^nT^nx) = (Q^ng,T^nx) \quad .$$

In view of Lemma 4.1 we may conclude that, for $R^n \sim (P \oplus Q^n)$,

$$\|\chi[(g,x)] - \chi[(g,Q^nx)]\|_{R^n} = \|\chi[(g,T^nx)] - \chi[(Q^ng,T^nx)]\|_{T^n}$$
$$\leq c \, \|(g - Q^ng,T^nx)\|_{T^n}$$
$$\leq c \, \|g - Q^ng\| \, \| \|T^nx\| \, \|_{T^n}$$

Since dim T^n is bounded, $\| \|T^nx\| \|_{T^n}$ is bounded uniformly with respect to n as $n \to \infty$. Also, $\|g - Q^ng\| \to 0$. Hence we obtain

$$\lim_{n\to\infty} \|\chi[(g,x)] - \chi[(g,Q^nx)]\|_{R^n} = 0 \qquad (x \in \mathcal{H}) \quad .$$

From remark (5.5), we may conclude, therefore, that $\chi[(g,\cdot)]$ is completely invariant as desired. Thus Lemma 7.2 and Theorem A are proved.

§ 8. Sufficient conditions for invariance.

In this section we return to the question raised in remark (5.5) on page 8. Though the problem of whether or not semi-invariance implies invariance has not been resolved, a few illuminating fragments will be discussed.

Let Π denote the set of semi-invariant proper functions and let Γ denote the set of "integrable" cylinder functions. Then Theorem A provides that $\Gamma \subset \Pi$. For any given functions f, let $(f+\Pi)$ denote the "coset" consisting of those functions which may be expressed as f plus a function of Π. Note that $\|g - f\|$ is defined if $g \in (f+\Pi)$.

Theorem: Suppose the function f is such that for any $\varepsilon > 0$ there exists a $g \in [\Gamma \cap (f+\overline{\Pi})]$ such that

$$\|g - f\| \leq \varepsilon \quad ,$$

then f is integrable and, moreover, invariant.

Proof: Given an $\varepsilon > 0$, the hypothesis provides a cylinder function g, whose base is in P such that $\|g - f\| \leq \varepsilon$. Furthermore, since $(g-f) \varepsilon \overline{\Pi}$ it follows that for any given complete basic system (Q^m)

$$\|Q^m g - Q^m f\|_R \leq 2\varepsilon \tag{8.1}$$

for all sufficiently large m. Let (P^n) be any other complete basic system. Since g is a cylinder function it is invariant by Theorem A. Hence, by the definition of invariance, we have

$$\|Q^m g - P^n g\|_R \leq \varepsilon \tag{8.2}$$

for all n and m sufficiently large. Also, analogously to (8.1), we have

$$\|P^n g - P^n f\|_R \leq \varepsilon \tag{8.3}$$

for all n sufficiently large. Finally, from (8.1), (8.2), and (8.3) we calculate that for all n and m sufficiently large

$$\|P^n f - Q^m f\|_R \leq 5\varepsilon$$

and hence f is invariant.

Corollary: If there exists a complete basic system (P^n) such that

(1) $(f - P^n f) \varepsilon \overline{\Pi}$ $(n = 1,2,\ldots)$

and (2) $f \varepsilon \mathcal{L}_a^{(P^n)}$

then f is invariant.

Proof: For any given n,

$$\|f - P^n f\| = \lim_{m\to\infty} \|P^m f - P^m P^n f\|_{\hat{R}}$$

$$= \lim_{m\to\infty} \|P^m f - P^n f\|_{(P^n)} .$$

Thus, since $f \in \mathcal{L}_a^{(P^n)}$, we get

$$\lim_{n\to\infty} \|f - P^n f\|_{\hat{R}} = 0 \tag{8.4}$$

From hypothesis (1) we have $g_n = P^n f \in [\Gamma \cap (f+\Pi)]$ and (8.4) then yields the remainder of the hypothesis of the theorem. Hence f is invariant.

§ 9. General Gaussian Measures.

In all of the above development the measures associated with the finite dimensional subspace P was given by

$$dm_P = \sqrt{2\pi}^{-K} \exp[-\tfrac{1}{2}\|Px\|^2]dx_1,\ldots dx_K$$

where K = dim P. All of the above discussion and results carry over without change if we use, instead, the measure,

$$d^\sigma m_P = \sqrt{\sigma/\pi}^{K} \exp[-\sigma\|Px\|^2]dx_1,\ldots dx_K \tag{9.1}$$

with $\sigma > 0$ fixed.

The question then arises as to how much carries over if we use "complex Gaussian measures"

$$d^z m_P = \sqrt{z/\pi}^{K} \exp[-z\|Px\|^2\| dx_1,\ldots dx_K \tag{9.2}$$

z fixed, $\sigma = \mathrm{Re}\ z > 0$. Though the formalisms of the basic set-up carry over to this case, some of the notions and results do not. In particular, it is here appropriate to define $f(x) \in \mathcal{L}_a^P$ for a cylinder function with base in P to mean that $|f(x)|^a$ is integrable over P with respect to the $d^\sigma m_P$ measure. Also, we define in this case

$$\|f\|_P^z = \left| \sqrt{z/\pi}^{K} \int_P |f(Px)|^a \exp[-z\|Px\|^2]dx_1\ldots dx_K \right|^{\frac{1}{2}}$$

where the z in the norm notation indicates the z used in the measure $d^z m_P$.

The main result which carries over to the case of these complex measures is essentially Theorem A, in the following form

Theorem: If $f(x)$ is a cylinder function with base in P such that $f \in \mathcal{L}_a^P$ then $f(x)$ is completely invariant.

Proof: Given a simply complete sequence of projections (Q^n), we wish to prove that

$$\lim_{m,n\to\infty} \|Q^m f - Q^n f\|_{\hat{R}}^z = 0 \quad .$$

To obtain this it clearly suffices to prove that

$$\lim_{m\to\infty} \|f - Q^m f\|_{\hat{R}}^z = 0 \quad . \tag{9.3}$$

As was done in §6, and §7, it is easily seen that for $T^m \sim (P \oplus Q^m P)$ we have

$$\|f - Q^m f\|_{\hat{R}}^z = \|(T^m f - T^m Q^m f)\|_{T^m}^z$$

$$\leq \left(\frac{|z|}{\sigma}\right)^{\left(\frac{\dim T^m}{2}\right)} \|T^m f - T^m Q^m f\|_{T^m} \quad .$$

Since $\dim T^m \leq 2 \dim P$ and f is completely invariant in the theory of the measures $d^\sigma m_p$, it follows that the right side of the above inequality $\to 0$ as $m \to \infty$. Hence (9.3) follows, which completes the proof of the theorem.

Chapter VI

A Property of Gaussian Measures

§ 1. Introduction.

In the development of a theory of integration over Hilbert space utilizing the Gaussian measures heavy use is made of the unitary invariance of such a measure over a finite dimensional Euclidean space. The use of the Gaussian measures is in fact dictated by a desire to have this property of the measure.

We now propose to prove that the property of unitary invariance characterizes the Gaussian measure among all normalized product measures. More precisely, we shall prove:

Theorem: Let E_n be n-dimensional Euclidean space, and for a given coordinate system, $x = (x_1,\ldots,x_n)$, consider the measure

$$\mu(dx) = dF_1(x_1) \ldots dF_n(x_n)$$

where the F_i $(i = 1,\ldots,n)$ are certain distribution functions. If we assume that $\mu(dx)$ remains invariant under all unitary transformations of E_n, then for some σ,

$$F_1(\xi) = F_2(\xi) = \ldots = F_n(\xi) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\xi} \exp\left[-\frac{u^2}{2\sigma^2}\right]du \ .$$

§ 2. Proof of the Theorem

Since $\mu(dx)$ must be invariant under a transposition of two of the coordinates it follows easily that all the $F_i(x)$ must be the same distribution function $F(x)$.

Thus it suffices to prove the theorem in the case n=2, in which case we denote the coordinates by x and y. Consider then, for any w, the set $A_w = \{(x,y):\ y \le w\}$. Given any pair of real numbers α,β such that $\alpha^2+\beta^2 = 1$, the unitary transformation

$$\begin{pmatrix} \beta & \alpha \\ -\alpha & \beta \end{pmatrix}$$

carries A_w into

$$A'_w = \{(x,y): \ \alpha x + \beta y \leq w\} \ .$$

Since A_w and A'_w must have the same measure we obtain

$$(2.1) \qquad F(w) = \iint_{\alpha x+\beta y \leq w} \alpha F(x)\beta F(y) \ .$$

Letting $\phi(t) = \int_{-\infty}^{\infty} e^{itx} dF(x)$, the characteristic function of F, (2.1) implies that

$$(2.2) \qquad \phi(t) = \phi(\alpha t)\phi(\beta t)$$

for all α,β such that $\alpha^2+\beta^2 = 1$. Taking $\alpha = -1$, $\beta = 0$ yields in particular

$$(2.3) \qquad \phi(t) = \phi(-t)$$

Taking $t \geq 0$ and setting $u = \alpha t$ and $v = \beta t$, (2.2) yields

$$\phi(u)\phi(v) = \phi(\sqrt{u^2+v^2})$$

or

$$\phi(\sqrt{u})\phi(\sqrt{v}) = \phi(\sqrt{u+v}) \ , \qquad u,v \geq 0 \ .$$

Setting $g(u) = \phi(\sqrt{u})$, we obtain

$$(2.4) \qquad g(u)g(v) = g(u+v) \ , \qquad u,v \geq 0 \ .$$

Since $\phi(u)$ is continuous, it follows that, for $u \geq 0$, $g(u)$ is a continuous solution of (2.4). Hence, as is well-known, for $u \geq 0$

$$g(u) = e^{-\lambda u}$$

and consequently

$$(2.5) \qquad \phi(u) = e^{-\lambda u^2}$$

Though (2.5) has been derived for $u \geq 0$, by (2.3) it extends immediately to all real u.

But $e^{-\lambda u^2}$ is precisely the characteristic function of a Gaussian distribution with mean 0 and variance $\sqrt{2\lambda}$. Since a characteristic function determines a distribution function uniquely at its points of continuity, it follows that F(x) must be this Gaussian distribution.

Chapter VII

The Integration of Polynomial Functionals

The simplest kind of function $f[\cdot]$, defined for elements x of a real Hilbert space $\mathcal{H}$, for which the integral (see Chap.I)

$$I(f) = \int^{\infty} f[x]\, e^{-\frac{1}{2}(x,x)} \prod_s \frac{\delta x_s}{\sqrt{2\pi}}$$

can be evaluated are polynomials.

We shall first define homogeneous polynomials of degree r - for short, "r-ics". To define such polynomials we may employ a representation of $\mathcal{H}$ by the set $\{\phi(\cdot)\}$ of the real-valued functions of finite norm where the unit form is given by

$$(x,x) = \|x\|^2 = \int \phi^2(s)\, \mu(ds) < \infty$$

where ϕ is the function representing the element x of $\mathcal{H}$. Let $\ell(s_1 \ldots s_r)$ be a square-integrable function of the r variables $s_1, \ldots, s_r$ so that

$$\int \ldots \int \ell^2(s_1 \ldots s_r)\mu(ds_1)\ldots\mu(ds_r) < \infty \ .$$

Then every $x \in \mathcal{H}$ is represented by a square-integrable function $\phi(\cdot)$ and, for every such x, the functional $\ell[x]$ given, in this representation, by the integral

$$\ell[x] = \int \ldots \int \ell(s_1 \ldots s_r)\phi(s_1)\ldots\phi(s_r)\mu(ds_1)\ldots\mu(ds_r)$$

is defined and, similarly for $x^{(1)}, \ldots, x^{(r)} \in \mathcal{H}$, the associated multilinear form

$$\ell[x^{(1)} \ldots x^{(r)}] = \int \ldots \int \ell(s_1 \ldots s_r)\phi^{(1)}(s_1)\ldots\phi^{(r)}(s_r)\mu(ds_1)\ldots\mu(ds_r)$$

is well-defined.

It is well-known - and easily proved - that the form $\ell[x]$ is independent of the particular representation of $\mathcal{H}$ by functions $\{\phi\}$. Thus, $\ell[x]$ may be regarded directly as a functional on $\mathcal{H}$ - a fact anticipated by our use of the notation "$\ell[x]$".

Similarly, $\ell[x^{(1)}\ldots x^{(r)}]$ may be regarded directly as a functional of the r elements of $\mathcal{R}$, $x^{(1)},\ldots,x^{(r)}$.

A polynomial $p_n[x]$ of degree $\leq n$ may now be defined as a sum,

$$p_n[x] = \sum_{r=1}^{n} \ell_r[x] \quad ,$$

of monomials of degree r (r-ics) where ℓ_o is simply a constant. Such polynomial functionals were called "regular" by Gateaux. Note that there exist non-regular polynomials as, for example, the unit form (x,x). We shall not be concerned here with the integration of such non-regular polynomials.

Instead of evaluating the integral I(p) directly we shall first evaluate the square-integral $I(p^2)$ and the integral I(qp) of the product of two such polynomials. The integral I(p) can then be obtained as I(qp) with $q \equiv 1$.

According to its definition (III), the integral I(f) is to be evaluated with the aid of a basic system (P) of projections P^σ.

We naturally employ a representation of the element x by a function $\phi(\cdot)$ such that the element $P^\sigma x$ is represented by a piecewise constant function

$$P^\sigma x \longleftrightarrow \phi^\sigma(s) = \begin{cases} \phi_\nu^\sigma & \text{for } s \in \mathcal{J}_\nu^\sigma \end{cases}$$

where $\mathcal{J}_\nu^\sigma$ is an interval of a partition $\mathcal{P}^\sigma$ and

$$\phi_\nu^\sigma = \int_{\mathcal{J}_\nu^\sigma} \phi(s)\mu(ds) / \Delta_\nu^\sigma$$

with

$$\Delta_\nu^\sigma = \int_{\mathcal{J}_\nu^\sigma} \mu(ds) = \mu(\mathcal{J}_\nu^\sigma)$$

Then f[·] is integrable relative to the basic system of projections (P) - or "(P)-integrable" for short - if

$$I^{(P)}\left(|P^\sigma f - P^\tau f|^2\right) \to 0 \quad \text{as } \sigma,\tau \to \infty \ .$$

The value of the integral $I^{(P)}(f^2)$ is now defined by

$$I^{(P)}(f^2) = \lim_{\sigma\to\infty} I^{(P)}(P^\sigma f^2) \quad .$$

If $f = \ell[\cdot]$ is an r-ic we have

$$P^\sigma f[x] = f[P^\sigma x] = \sum_{\nu_1\cdots\nu_r} \ell^\sigma_{\nu_1\cdots\nu_r}\, \phi^\sigma_{\nu_1}\cdots\phi^\sigma_{\nu_r}\, \Delta^\sigma_{\nu_1}\cdots\Delta^\sigma_{\nu_r}$$

with

$$\ell^\sigma_{\nu_1\cdots\nu_r} = \int_{\mathcal{I}^\sigma_{\nu_1}}\cdots\int_{\mathcal{I}^\sigma_{\nu_r}} \ell(s_1\cdots s_r)\mu(ds_1)\cdots\mu(ds_r) / \Delta^\sigma_{\nu_1}\cdots\Delta^\sigma_{\nu_r}$$

Let ρ^σ be a refinement of ρ^τ. Then the functional $f[P^\tau x]$, which is a cylinder functional with respect to P^τ, is also a cylinder functional with respect to P^σ. Consequently, the integral I^{P^σ} is defined for $|P^\sigma f - P^\tau f|^2$, and we have

$$I^{(P)}(|P^\sigma f - P^\tau f|^2) = I^{P^\sigma}(|P^\sigma f - P^\tau f|^2) \quad .$$

We confine ourselves at present to evaluating the integral of the square of a form, $\ell[\phi] = \ell_2[\phi]$, of second degree.

First we assume $\ell[\phi]$ to be a cylinder functional with respect to P;

$$\ell[\phi] = \sum_{\nu,\mu} \ell_{\nu\mu}\phi_\nu\phi_\mu\Delta_\nu\Delta_\mu \quad .$$

Then we find

$$I^{(P)}(\ell^2) = \int\cdots\int \left\{\sum_{\nu,\mu} \nu_\mu\phi_\nu\phi_\mu\Delta_\nu\Delta_\mu\right\}^2 e^{-\frac{1}{2}\sum_\nu\phi^2_\nu\Delta_\nu} \Pi_\nu \sqrt{\frac{\Delta_\nu}{2\pi}}\, d\phi_\nu \quad .$$

Using the formula

$$\int \xi^{2k}\, e^{-\frac{1}{2}a\xi^2} \sqrt{\frac{a}{2\pi}}\, d\xi = \frac{(2k)!}{2^k\, k!} \quad ,$$

valid for $k = 0,1,2,..$, we find

$$I^{(P)}(\ell^2) = \sum_{\nu\neq\mu} \ell_{\nu\nu}\ell_{\mu\mu}\Delta_\nu\Delta_\mu + 2\sum_{\nu\neq\mu} \ell^2_{\nu\mu}\Delta_\nu\Delta_\mu + 3\sum_\nu \ell^2_{\nu\nu}\Delta^2_\nu)$$

$$= (\sum_\nu \ell_{\nu\nu}\Delta_\nu)^2 + 2\sum_{\nu,\mu} \ell^2_{\nu,\mu}\Delta_\nu\Delta_\mu \quad .$$

This result may be written in the form

$$(*)\quad I^{(P)}(\ell^2) = 2\iint \ell^2(s_1,s_2)\mu(ds_1)\mu(ds_2) + [\int \ell(s,s)\mu(ds)]^2 \ .$$

Naturally, one expects that the same formula (*) holds for forms $\ell[x]$ of the second degree which are <u>not</u> cylinder functionals.

It is clear that formula (15) may be applied as it stands to the cylinder functional $|P^\tau\ell - P^\sigma\ell|^2$. Hence we have

$$I^{(P)}(|P^\tau\ell - P^\sigma\ell|^2) = 2\iint [(P^\tau\ell - P^\sigma\ell)^2](s_1,s_2)\mu(ds_1)\mu(ds_2) + [\int [(P^\tau\ell - P^\sigma\ell)^2](s,s)\mu(ds)]^2 \ .$$

From the square integrability of the kernel of the form ℓ, it is clear that

$$\iint [(P^\tau\ell - P^\sigma\ell)^2](s_1,s_2)\mu(ds_1)\mu(ds_2) \to 0 \qquad \text{as } \sigma,\tau \to \infty \ .$$

Suppose that, in addition, the relation

$$(**)\quad \int [(P^\tau\ell - P^\sigma\ell)^2](s,s)\mu(ds) \to 0 \qquad \text{as } \sigma,\tau \to \infty$$

is valid. Then we can conclude that

$$I^{(P)}(|P^\sigma\ell - P^\tau\ell|^2) \to 0 \qquad \text{as } \sigma,\tau \to \infty$$

and, hence, that the form $\ell[\cdot]$ is square integrable.

Relation (**) implies that the sequence $\int P^\sigma\ell(s,s)\mu(ds)$ has a limit. Naturally, we shall employ the notation "$\int \ell(s,s)\mu(ds)$" for this limit;

$$\int [P^\sigma\ell](s,s)\mu(ds) \to \int \ell(s,s)\mu(ds) \ .$$

Formula (*) then would hold for the form $\ell[\cdot]$.

Thus we are led to investigate the conditions under which relation (**) holds. Since the integral $\int \ell(s,s)\mu(ds)$ is called the "trace" of the form $\ell[\cdot]$,

$$\mathrm{Tr}\ \ell = \int \ell(s,s)\mu(ds) \ ,$$

we may also ask: does the form ℓ possess a trace? If so the form will be called "traceable" (*).

Obviously, there are forms which do not possess a trace, since there are functions $\ell(s_1,s_2)$ for which $\iint \ell^2(s_1,s_2)\mu(ds_1)\mu(ds_2)$ is finite but $\int \ell(s,s)\mu(ds)$ is not.

However, even if the trace exists something may happen which we should like to exclude. Namely, it may happen that the trace depends on the choice of the basic system (P). In this case the value of the integral $I(\ell^2)$ would also be dependent on this choice. Suppose the function $\ell(s,s)$ is singular in such a way that the integral is only conditionally convergent. Then we know that the value of the integral may be altered by appropriately rearranging the sequence $\{P^\sigma\}$. In order to exclude this occurrence we require that the function $\ell(s,s)$ be absolutely integrable,

$$\int |\ell(s,s)|\mu(ds) < \infty \quad .$$

If this holds we shall say the form $\ell[\cdot]$ is "absolutely traceable".

This requirement, however, is still insufficient to ensure the independence of the trace from the choice of the representation. We consider the following counter-example.

Let each element x of the space be represented by a sequence,

$$x \longleftrightarrow \{y_1,z_1,y_2,z_2\cdots\}$$

with

$$\|x\| = \sum_{\nu=1}^{\infty} [y_\nu^2 + z_\nu^2] \quad ,$$

and consider the form

(*) The term "betraced" has been suggested as being more appropriate than "traceable". This would be so if one imagined the trace as something put on the form. We prefer to think of the trace as being what is "left" after a certain operation has been carried out.

$$\mathcal{l}[x] = \sum_{\nu=1}^{\infty} [\varepsilon_\nu(y_\nu^2 + z_\nu^2) + 2y_\nu z_\nu]$$

where $\varepsilon_\nu > 0$ for $\nu = 1,2\ldots$ and $\sum_{\nu=1}^{\infty} \varepsilon_\nu < \infty$. Evidently then,

$$\mathrm{Tr}\ \mathcal{l} = 2 \sum_{\nu=1}^{\infty} \varepsilon_\nu < \infty \ .$$

We may also write $\mathcal{l}[x]$ in the form

$$\mathcal{l}[x] = \sum_{\nu=1}^{\infty} \left[(1 + \varepsilon_\nu) \frac{(y_\nu+z_\nu)^2}{2} - (1 - \varepsilon_\nu) \frac{(y_\nu-z_\nu)^2}{2} \right]$$

and, introducing $\eta_\nu = \frac{y_\nu+z_\nu}{\sqrt{2}}$ and $\zeta_\nu = \frac{y_\nu-z_\nu}{\sqrt{2}}$ as the components of a new representation of the element x, we may so choose the order of these new components that the trace of $\mathcal{l}$ is arbitrarily close to the number

$$a(\lambda,\mu) = \left[\sum_{\nu=1}^{\lambda} (1 + \varepsilon_\nu) - \sum_{\nu=1}^{\mu} (1 - \varepsilon_\nu) \right]$$

with arbitrary λ and μ. Thus, here the trace <u>does</u> depend on the coordinate system. In other words, the trace does not converge absolutely with respect to the new coordinate system.

If the trace of a form is independent of the choice of the representation it will be called "invariantly traceable". If the form $\mathcal{l}$ has this property the evaluation (15) of $I(\mathcal{l}^2)$ holds independently of the representation.

A <u>necessary and sufficient condition</u> for <u>invariant traceability</u> can be given: The form $\mathcal{l}$ can be written as the difference

$$\mathcal{l}[x] = \mathcal{l}^+[x] - \mathcal{l}^-[x]$$

of two traceable positive bounded forms, $\mathcal{l}^+$ and $\mathcal{l}^-$.

<u>Proof</u>: It is well known(*) that the trace of a positive form is invariant. The same is true, therefore, of the

---

(*) Von Neumann, Math. Found. of Quantum Mechanics, Chapter II.

difference of two such forms. Conversely, let the form $\ell[\cdot]$ be invariantly traceable; then this form, being bounded, can be diagonalized in the form

$$\ell[x] = \int \lambda(s)\eta^2(s)\mu(ds)$$

where $\eta(\cdot)$ is the representer of x in another representation. From the invariant traceability of one concludes that in this representation, too, the absolute trace exists:

$$\int |\lambda(s)|\mu(ds) < \infty \quad .$$

Since the trace exists also with respect to the diagonal representation we have

$$\mathrm{Tr}\,\ell = \int \lambda(s)\mu(ds)$$

It is then clear that

$$\int_{\lambda>0} \lambda(s)\mu(ds) < \infty \quad \text{and} \quad \int_{\lambda<0} -\lambda(s)\mu(ds) < \infty$$

from which it follows that ℓ can be written as

$$\ell[x] = \int \lambda^+(s)\eta^2(s)\mu(ds) - \int \lambda^-(s)\eta^2(s)\mu(ds)$$

where

$$\lambda^{\pm}(s) = \tfrac{1}{2}\,|\lambda(s) \pm |\lambda(s)||\quad .$$

It is immediately clear that the terms $\ell^{\pm}$ are bounded and positive.

It will be necessary in the following to introduce the notion of the trace of a form of higher than second degree. To the m-ic form $\ell[\cdot] = \ell_m[\cdot]$ we assign the muïtilinear form

$$\ell[x^{(1)},\ldots x^{(m-2)},y^{(1)},y^{(2)}] \quad .$$

$$= \int\ldots\int \ell(s_1\ldots s_m)\phi^{(1)}(s_1)\ldots\phi^{(m-2)}(s_{m-2})\psi^{(1)}(s_{m-1})\psi^{(2)}(s_m)\mu(ds_1)\ldots\mu(ds_m)$$

where $x^{(i)} \longleftrightarrow \phi^{(i)}(\cdot)$ and $y^{(j)} \longleftrightarrow \psi^{(j)}(\cdot)$. We note that $\ell[x^{(1)}\ldots y^{(1)},y^{(2)}]$ is bilinear in $y^{(1)}$ and $y^{(2)}$ and denote by

Tr ℓ the trace of this form with respect to $y^{(1)}$ and $y^{(2)}$,

$$(\mathrm{Tr}\,\ell)[x^{(1)},\dots x^{(m-2)}]$$
$$= \int^{m-1} \ell(s_1 \dots s_{m-2}, s, s)\phi^{(1)}(s_1)\dots\phi^{(m-2)}(s_{m-2})\mu(ds_1)\dots\mu(ds_{m-2})\mu(ds)$$

which is evidently itself an (m-2)-ic form. If this form is independent of the representation we call $\ell[\cdot]$ invariantly traceable; if all the forms,- ℓ, Tr ℓ, $(\mathrm{Tr})^2\ell,\dots(\mathrm{Tr})^{[m/2]-1}\ell$ - are invariantly traceable we call ℓ "totally invariantly traceable". Incidentally, any product of bounded linear forms is totally invariantly traceable.

A simple formal computation shows that the integral of a form ℓ_{2k} of degree 2k is given by the formula

$$I(\ell_{2k}) = (2k)!\ [(\mathrm{Tr})^k \ell]/2^k k! \quad .$$

If, in particular, we take ℓ_{2m} as the square of a form ℓ_m of degree m we get, after some computation, the formula

$$I(\ell_m^2) = \sum_{k=0}^{[m/2]} C_{m,k} \int^{m-2k} C_{m,k} \int^{m-2k} A(t_1 \dots t_{m-2k}!)\mu(dt_1)\dots\mu(dt_{m-2k})$$

where

$$A(t_1 \dots t_{m-2k} = \int^{k} (s_1, s_1, \dots s_k, s_k, t_1 \dots t_{m-2k})\mu(ds_1)\dots\mu(ds_k)$$

in which the $C_{m,k}$ are positive coefficients among which, in particular, $C_{m,0} = m!$.

In many respects it is more appropriate to consider "Hermite polynomials" in place of monomial forms. In order to define Hermite polynomial functionals it is convenient to introduce the symbolic operator $\delta/\mu(ds)\delta\phi(s)$ which transforms the m-ic form

$$\ell_m[x] = \int \dots \int \ell(s_1 \dots s_m)\phi(s_1)\dots\phi(s_m)\mu(ds_1)\dots\mu(ds_m)$$

into the (m-1)-ic form, with parameter s,

$$\frac{\delta}{\mu(ds)\delta\phi(s)}\ell_m[x] = m\int \dots \int \ell(s_1 \dots s_{m-1}, s)\phi(s_1)\dots\phi(s_{m-1})\mu(ds_1)\dots\mu(ds_{m-1}) .$$

The operator symbolically written as

$$L = \int \left[\frac{\delta}{\mu(ds)\delta\phi(s)} \right]^2 \mu(ds) \quad ,$$

which is the analogue of the Laplacian, then yields a multiple of the trace:

$$L\, \ell_k = k(k-1)[\mathrm{Tr}\ \ell_k] \quad .$$

We use, too, the symbolic expression $\ell[x - \frac{\delta}{\delta x}]$ for the form obtained by substituting

$$\left[\phi(s_\nu) - \frac{\delta}{\mu(ds_\nu)\delta\phi(s_\nu)} \right]$$

for each $\phi(s_\nu)$ in the integral expression for $\ell[x]$ and then substituting zero for the operator $\delta/\mu(ds)\delta\phi(s)$ if it does not precede a function $\phi(\cdot)$. Thus in particular,

$$\ell_2[x-\frac{\delta}{\delta x}] = \int\int \ell(s,t)\left[\phi(s)-\frac{\delta}{\mu(ds)\delta\phi(s)}\right]\left[\phi(t)-\frac{\delta}{\mu(dt)\delta\phi(t)}\right]\mu(ds)\mu(dt)$$
$$= \ell_2[x] - \int \ell(s,s)\mu(ds) = \ell_2[x] - \mathrm{Tr}\ \ell_2 \quad .$$

Now, to each m-ic ℓ_m we assign an "Hermite polynomial" functional $H\ell_m$ by the prescription

$$H\ell_m[x] = \ell[x - \frac{\delta}{\delta x}] \quad .$$

Thus, for example, we have

$$H\ell_0[x] = \ell_0 \quad ,$$
$$H\ell_1[x] = \ell_1[x] \quad , \text{ and}$$
$$H\ell_2[x] = \ell_2[x] - \mathrm{Tr}\ \ell_2 \quad .$$

In general

$$H\ell_n[x] = \sum_{k=0}^{[n/2]} \frac{(-1)^k n!(\mathrm{Tr})^k \ell_n[x]}{2^k k!(n-2k+1)!} \quad .$$

Assuming that both the homogeneous functionals $\hat{\ell}$ and $\tilde{\ell}$ are totally invariantly traceable we can evaluate the integral of

their product. Let $\hat{\ell}$ and $\tilde{\ell}$ have degrees m and n respectively. Then, abbreviating $\delta/\mu(ds)\delta\phi(s)$ to $\bar{\delta}s$, we have

$$I(H\hat{\ell}H\tilde{\ell}) = \int^{\infty}\left\{\int^{m}\hat{\ell}(s_1\cdots s_m)[\phi(s_1)-\bar{\delta}s_1]\cdots[\phi(s_m)-\bar{\delta}s_m]\mu(ds_1)\cdots\mu(ds_m)\right\}\cdot$$

$$\cdot\left\{\int^{n}\tilde{\ell}(s_{m+1}\cdots s_{m+n})[\phi(s_{m+1})-\bar{\delta}s_{m+1}]\cdots[\phi(s_{m+n})-\bar{\delta}s_{m+n}]\cdot\right.$$

$$\left.\mu(ds_{m+1})\cdots\mu(ds_{m+n})\right\}\cdot e^{-\frac{1}{2}\int\phi^2(s)\mu(ds)}\prod_s\sqrt{\frac{\mu(ds)}{2\pi}}\,\delta\phi(s)$$

$$=\int^{m+n}\hat{\ell}(s_1\cdots s_m)\tilde{\ell}(s_{m+1}\cdots s_{m+n})K(s_1\cdots s_{m+n})\mu(ds_1)\cdots\mu(ds_{m+n})$$

where

$$K(s_1\cdots s_{m+n}) = \int^{\infty}\left\{[\phi(s_1)-\bar{\delta}s_1]\cdots\right\}\left\{[\phi(s_{m+1})-\bar{\delta}s_{m+1}]\cdots\right\}$$

$$\cdot e^{-\frac{1}{2}\int\phi^2(s)\mu(ds)}\prod_s\sqrt{\frac{\mu(ds)}{2\pi}}\,\delta\phi(s)\quad .$$

Observing that $K(s_1\cdots s_{m+n})$ is the integral of a functional which, as it depends on the values of $\phi(\cdot)$ only at the points $s_1\cdots s_{m+n}$, is a cylinder functional, we may write K as a finite integral;

$$K(s_1\cdots s_{m+n}) = \int^{m+n}\left\{[\phi(s_1)-\bar{\delta}s_1]\cdots\right\}\left\{[\phi(s_{m+1})-\bar{\delta}s_{m+1}]\right.$$

$$\left.\cdot\exp\left[-\sum_k\phi^2(s_k)\mu(ds_k)\right]\prod_k\sqrt{\frac{\mu(ds_k)}{2\pi}}\,\delta\phi(s_k)\right.\quad .$$

Observing farther that each $\phi(s_k)$ is, here, simply an independent variable (noting that if, e.g., $s_1=s_2$ then $\phi(s_1)$ and $\phi(s_2)$ are, of course, the same), we replace $\phi(s_k)$ by ξ_k (keeping necessary identifications in mind) and similarly replace $\bar{\delta}s_k = \delta/\mu(ds_k)\delta\phi(s_k)$ by $d/d\xi_k$. We recognize that

$$\left\{[\xi_1-\frac{d}{d\xi_1}]\cdots[\xi_m-\frac{d}{d\xi_m}]\right\}\text{ and }\left\{[\xi_{m+1}-\frac{d}{d\xi_{m+1}}]\cdots[\xi_{m+n}-\frac{d}{d\xi_{m+n}}]\right\}$$

are m-th and n-th degree multi-dimensional Hermite polynomials respectively (which we denote by H_m and H_n). Then

$$K(s_1 \cdots s_{m+n}) =$$

$$\int H_m(\xi_1 \cdots \xi_m) H_n(\xi_{m+1} \cdots \xi_{m+n}) e^{-\sum_k \xi_k^2 \mu(ds_k)} \prod_k \sqrt{\frac{\mu(ds_k)}{2\pi}}\, d\xi_k \quad .$$

From the theory of multi-dimensional Hermite polynomials we conclude that K will vanish unless m = n and each variable appearing in H_m may be paired with (i.e. - is identical to) a variable in H_n and vice-versa as, for example, $s_1 = s_{m+1}, \ldots, s_m = s_{m+n} (= s_{2m})$. Then

$$K(s_1 \cdots s_m, s_{i_1} \cdots s_{i_m}) = \int^m H_m^2(\xi_1 \cdots \xi_m) e^{-\frac{1}{2}\sum_k \xi_k^2 \mu(ds_k)} \prod_k \sqrt{\frac{\mu(ds_k)}{2\pi}}\, d\xi_k$$

$$= 1$$

where $\{i_1 \ldots i_m\}$ is any permutation of $\{1 \ldots m\}$. As there are m! such permutations (and as we have assumed the kernels $\hat{\ell}(s_1 \cdots s_m)$ and $\tilde{\ell}(s_1 \cdots s_m)$ to be symmetric in all variables), we have

$$I(H\hat{\ell} H\tilde{\ell}) = m! \int^m \hat{\ell}(s_1 \cdots s_m) \tilde{\ell}(s_1 \cdots s_m) \mu(ds_1) \cdots \mu(ds_m) \quad .$$

In other words, two Hermite polynomial functionals, $H\hat{\ell}[x]$ and $H\tilde{\ell}[x]$ are orthogonal to each other with respect to the unit form given by the integral $I(\cdot)$ if their degrees are unequal or, even if the degrees are the same, if the kernels, $\hat{\ell}(\ldots)$ and $\tilde{\ell}(\ldots)$, are orthogonal to each other. Since this orthogonality property characterizes the Hermite polynomials of one variable or any finite number of variables (for which, also, an analogue of the formula, $H\ell[x] = \ell\left[x - \frac{\delta}{\delta x}\right]$, holds), the name "Hermite polynomials" for the functionals here introduced appears justified. (For a somewhat different description of these polynomials see "Mathematical Aspects of the Quantum Theory of Fields" by K. O. Friedrichs; Interscience, 1953, p.52 et seq. and the references given there.)

The advantage of working with Hermite polynomials is this; Suppose two functionals, $\hat{f}[\cdot]$ and $\tilde{f}[\cdot]$, admit expansions of the form

$$\hat{f}[x] = \sum_{m=0}^{\infty} \frac{1}{m!} H\hat{\ell}_m[x] \qquad \text{and} \qquad \tilde{f}[x] = \sum_{m=0}^{\infty} \frac{1}{m!} H\tilde{\ell}_m[x] \quad .$$

(We note that in the expansion the kernels of the forms play the part of expansion coefficients.) Then the integral of the product is given by

$$I(\hat{f}\tilde{f}) = \sum_{m=0}^{\infty} \frac{1}{m!} \int^m \hat{\ell}_m(s_1 \cdots s_m)\tilde{\ell}_m(s_1 \cdots s_m)\mu(ds_1)\cdots\mu(ds_m) \quad .$$

In fact, this formula could be used as a definition of the integral $I(\cdot)$ which is a priori independent of the representation of the element of $\mathcal{H}$ by functions since the products $(\hat{\ell}_m\tilde{\ell}_m)$ are (see K.O. Friedrichs, loc. cit. and references there).

There are other uses of the expansion with respect to Hermite polynomials. For example, a Fourier transformation for functionals can be described very easily with the aid of such an expansion. The Fourier transformation should be applied, not to functionals which are square integrable with respect to the integral $I(\cdot)$, but to functionals $f[\cdot]$ which are square integrable with respect to the translation invariant integral $J(\cdot)$ (see I-9). We may also say the Fourier transformation should be applied to functions $\exp[-\frac{1}{4}(x,x)]f[x]$ where $f[x]$ is square integrable with respect to $I(\cdot)$. Here we have assumed the exponent entering the Gaussian measure is $-\frac{1}{2}(x,x)$. Had it been $-a(x,x)$ we would consider $\hat{f}[x] = \exp[-\frac{a}{2}(x,x)]f[x]$ as the square integrability of f with respect to $I(\cdot)$ would, then, ensure the square integrability of $\hat{f}$ with respect to $J(\cdot)$.

Let $\phi(\cdot)$ represent the independent variable x of $f[x]$ and, similarly, let $\psi(\cdot)$ represent the y in the transform, $g[y] = (Gf)[y]$. We write the inner product for the representations as;

$$(x,y) = \phi\cdot\psi = \int \phi(s)\psi(s)\mu(ds) \quad , \quad \text{etc.}$$

We also write, somewhat inconsistently, $f = f[\phi]$ and $g = g[\psi]$. We then write the transformation in the form

$$(Gf)[\psi] = e^{\frac{1}{4}\psi\cdot\psi} \int^{\infty} e^{\frac{1}{2}i\psi\cdot\phi} f[\phi] e^{-\frac{1}{4}\phi\cdot\phi} \prod_s \sqrt{\frac{\mu(ds)}{4\pi}}\, \delta\phi(s) \quad .$$

We have, here, chosen the exponent $\frac{i}{2}\psi\cdot\phi$ rather than $i\psi\cdot\phi$ because the subsequent formulas become more elegant. [Gl = 1.]

The integral on the right-hand side may be regarded as an integral of the same type as $I(\cdot)$ but with $a = \frac{1}{4}$ instead of $a = \frac{1}{2}$. For this reason it was necessary to use the factor $\sqrt{\frac{\mu(ds)}{4\pi}}$ instead of $\sqrt{\frac{\mu(ds)}{2\pi}}$ with the differential. Only for rather special functionals f can one expect this integral to exist. Still the transformation can be given a meaning for all square integrable elements and for the ideal elements they generate.

To indicate how this can be done we assume that the functional $f[\cdot]$ admits the expansion with respect to Hermite polynomials,

$$f[\phi] = \sum_{m=0}^{\infty} \frac{1}{m!} H\ell_m[\phi] \quad .$$

Before defining its transform,

$$g[\psi] = (Gf)[\psi] \quad ,$$

we assign a transform to the Hermite polynomials $H\ell_m$. In doing this we guide ourselves by the formal expression given above. Accordingly, we set

$$(GH\ell_m)[\psi] = e^{\frac{1}{4}\psi\cdot\psi} \int^{\infty} e^{\frac{i}{2}\psi\cdot\phi}\, e^{-\frac{1}{4}\phi\cdot\phi}\, \ell_m[\phi(s) - \bar{\delta}s] \prod_s \sqrt{\frac{\mu(ds)}{4\pi}}\, \delta\phi(s).$$

Now we obviously have the identity

$$\left[\frac{1}{2}\phi(s) - \frac{\delta}{\mu(ds)\delta\phi(s)}\right] e^{-\frac{1}{4}\phi\cdot\phi} q[\phi] = e^{-\frac{1}{4}\phi\cdot\phi} \left[\phi(s) - \frac{\delta}{\mu(ds)\delta\phi(s)}\right] q[\phi]$$

and, more generally,

$$\left[\tfrac{1}{2}\psi(s) - i\frac{\delta}{\mu(ds)\delta\psi(s)} - \frac{\delta}{\mu(ds)\delta\phi(s)}\right]e^{-\frac{1}{4}\phi\cdot\phi+\frac{1}{2}\psi\cdot\phi}q[\phi]$$

$$= e^{-\frac{1}{4}\phi\cdot\phi+\frac{1}{2}\phi\cdot\phi}\left[\phi(s) - \frac{\delta}{\mu(ds)\delta\phi(s)}\right]q[\phi]$$

for any $q[\phi]$. Therefore we may write

$$(GH\ell_m)[\psi] = e^{-\frac{1}{4}\psi\cdot\psi}\int^{\infty}\ell_m\left[\tfrac{i}{2}\psi - \frac{\delta}{\mu(\cdot)\delta\psi} - \frac{\delta}{u(\cdot)\delta\phi}\right]$$

$$\cdot e^{\frac{1}{2}\psi\cdot\phi-\frac{1}{4}\phi\cdot\phi}\prod_s\sqrt{\frac{\mu(ds)}{4\pi}}\,\delta\phi(s) \quad .$$

Now, formally, for any $q(\phi,\psi)$ which vanishes as $\phi(s) \to \pm\infty$,

$$\int^{\infty}\frac{\delta}{\mu(ds)\delta\phi(s)}\,q[\phi,\psi]\prod_s\delta\phi(s) = 0 \quad .$$

Employing this formula m times we find that the term $\frac{\delta}{\mu(\cdot)\delta\phi}$ in the argument does not alter the result. Consequently we may write

$$(GH\ell_m)[\psi] = e^{\frac{1}{4}\psi\cdot\psi}\ell_m\left[\tfrac{i}{2}\psi - i\frac{\delta}{\mu(\cdot)\delta\psi}\right]\int^{\infty} e^{\frac{1}{2}\psi\cdot\phi-\frac{1}{4}\phi\cdot\phi}\prod_s\sqrt{\frac{\mu(ds)}{4\pi}}\,\delta\phi(s)$$

$$= e^{\frac{1}{4}\psi\cdot\psi}\,\ell_m\left[\tfrac{i}{2}\psi - i\frac{\delta}{\mu(\cdot)\delta\psi}\right]e^{-\frac{1}{4}\psi\cdot\psi}$$

$$= \ell_m\left[i\psi - i\frac{\delta}{\mu(\cdot)\delta\psi}\right]1 \qquad \text{by the identity on (VII-13)}$$

and, finally,

$$(GH\ell_m)[\psi] = i^m\,H\ell_m[\psi] \quad .$$

We may now employ this simple formula to define the Fourier transformations for any functional f possessing an expansion in Hermite polynomials. I.e., if

$$f[x] = \sum_{m=0}^{\infty} \frac{1}{m!} H\ell_m[x]$$

we define

$$g[y] = (Gf)[y] = \sum_{m=0}^{\infty} \frac{i^m}{m!} H\ell_m[x] \quad .$$

More precisely we should say in this case that

$$e^{-\frac{1}{4}(x,x)} \sum_{m=0}^{\infty} \frac{i^m}{m!} H\ell_m[x]$$

is the Fourier transform of

$$e^{-\frac{1}{4}(x,x)} f[x] \quad .$$

Chapter VIII

Integration over Hilbert Space and Related Norms

It has been seen in Chapter V that the system of definitions

$$dm_P = \sqrt{2\pi}^{-K} \exp[-\tfrac{1}{2}||Px||^2]d_P x \tag{3.1}$$

defines an integration space over Hilbert space $\mathfrak{H}$, in the sense of Chapter IV. Here P is any orthogonal projection onto a K-dimensional subspace of $\mathfrak{H}$. Thus, in terms of the theory developed in Chapter IV, we may define the Gaussian Integral.

$$I_G(f) = (2\pi)^{-K}\int f(Px)\exp[-\tfrac{1}{2}||Px||^2]d_P x$$

$$= \int_{\mathfrak{H}} f(x)\,\mu_G(x) = (2\pi)^{\frac{-\infty}{2}}\int_{\mathfrak{H}} f(x)e^{-1/2||x||^2}dx \tag{8.1}$$

of an arbitrary cylinder function such that $f(x)=f(Px)$; the last two forms are intended merely as suggestive ways of writing $I_G(f)$ as a formal integral, and are to be taken to be equal to $I_G(f)$ by definition.

Now at this stage, everything is evidently invariant relative to rotations in Hilbert space. More precisely let $\mathfrak{P}$ denote the set of all orthogonal projections P of $\mathfrak{H}$ onto one of its finite dimensional subspaces. Then, if U is a unitary transformation of $\mathfrak{H}$ onto itself, and $P \in \mathfrak{P}$, then $U^{-1}PU = Q$ still belongs to $\mathfrak{P}$. If f is a cylinder function defined on such that $f(x)=f(Px)$, then the function g defined by $g(x)=f(Ux)$, is also a cylinder function defined on $\mathfrak{H}$ and $g(Qx)=g(x)$; moreover, $I_G(f) = I_G(g)$; i.e.

$$\int_{\mathfrak{H}} f(Ux)\mu_G(dx) = \int_{\mathfrak{H}} f(x)\mu_G(dx). \tag{8.2}$$

The Gaussian integral of cylinder functions over Hilbert space has in addition the following evident properties:

$$\text{(a)}\quad \text{If } f \geq 0 \text{ then } I_G(f) \geq 0$$

(8.3) $$\text{(b)}\quad I_G(\alpha f+\beta g) = \alpha I_G(f) + \beta I_G(g)$$

$$\text{(c)}\quad I_G(|f(\cdot)|) \geq |I_G(f)|$$

all valid if f and g are integrable cylinder functions.

The problem now is to extend the range of definition of the Gaussian integral I_G from the relatively restricted class of integrable cylinder functions to more general classes of functions defined on Hilbert space. In chapters IV and V this was done by first extending the integral to a family of "ideal functions", and then by showing how a certain calss of functions actually defined on $\mathfrak{H}$ could be identified in various stronger or weaker senses with these "ideal functions". In the present chapter, an alternate attack on the problem of extending the Gaussian integral, which dispenses with the machinery of "ideal functions", will be outlined.

The general idea of this attack will be as follows: we will define a finite or infinite non-negative "norm" $||f||$ on the linear space of <u>all</u> functions f defined on $\mathfrak{H}$. The norm will satisfy $||af|| = |a|\,||f||$ for each scaler a, and $||f+g|| \leq ||f||+||g||$. Consequently, the set of functions with finite norms will form a linear subspace of the space of all functions defined on $\mathfrak{H}$; this subspace may conveniently be called the space of all <u>normable functions</u> defined on $\mathfrak{H}$. The norm will also be required to satisfy

(8.4)
$$\text{(a)}\quad \text{every integrable cylinder function } f \text{ is normable}$$

$$\text{(b)}\quad |I_G(f)| \leq ||f|| \text{ for each integrable cylinder function } f.$$

Suppose that the set of all integrable cylinder functions be denoted by the letter $\mathfrak{F}_0$, and the space of all normable functions by the letter F. Then, by (8.4)(b), the linear mapping

$I_G: f \longrightarrow I_G(f)$ is uniformly continuous on the subspace $\mathcal{L}_0$ of F, and hence has a unique continuous extension to the closure $\mathcal{L}$ of $\mathcal{L}_0$ in the topology of F determined by the norm. This closure $\mathcal{L}$ will be called space of integrable functions; the unique continuous extension of I_G from $\mathcal{L}_0$ to $\mathcal{L}$, which we shall continue to denote by the symbol I_G, will define the Gaussian integral $I_G(f)$ of an arbitrary function f in $\mathcal{L}$.

It should be remarked that we have not ruled out the possibility that $||f|| = 0$ even though $f \neq 0$. This slightly anomalous possibility causes no trouble however; in the familiar manner of integration theory, we simply call a function f such $||f|| = 0$ a null function, and identify any two functions g and h whose difference is a null function. Since $||g-h|| = 0$ implies $||ag-ah|| = 0$, ag and ah are identified if h and g are identified. Since $|\,||g|| - ||h||\,| \leq ||g-h||$, $||g|| = ||h||$ if g and h are identified. Let g and h be identified, and let $g \in \mathcal{L}$, so that there exists a sequence $g_n \in \mathcal{L}_0$ such that $||g_n-g|| \longrightarrow 0$, so that $I_G(g) = \lim_{n\longrightarrow\infty} I_G(g_n)$ by definition. Then

$||g_n-h|| \leq ||g_n-g|| + ||g-h|| = ||g_n-g||$, so that $h \in \mathcal{L}$ and

$I_G(h) = \lim_{n\longrightarrow\infty} I_G(g_n) = I_G(g)$. This shows that if g and h are identified, and $g \in \mathcal{L}$, then $h \in \mathcal{L}$ and the integrals $I_G(g)$ and $I_G(h)$ are equal. In this way we see that the identification of functions whose difference is a null-function is as harmless in the present theory as in ordinary integration theory. Similar elementary reasoning, making use of the triangle inequality, of (8.4), and of the definition of $\mathcal{L}$ as the closure of $\mathcal{L}_0$, will readily establish the following general principles.

(a) The set $\mathcal{L}$ of all integrable functions is a linear space containing the linear space of all integrable cylinder functions.

(8.5) (b) The Gaussian integral I_G is a linear functional defined on $\mathcal{L}$.

(c) $|I_G(f)| \leq ||f||$ for $f \in \mathcal{L}$

Suppose in addition that the norm $||f||$ is known to be unitary-invariant in the sense that $||f|| = ||g||$ if $g(x) = f(Ux)$, where U is a unitary mapping of Hilbert space h onto itself. Then it is evident that similar elementary reasonings will show that

(8.6) If $f \in \mathcal{F}$, and $g(x) = f(Ux)$, where U is a unitary mapping of h onto itself, then $g \in \mathcal{F}$, and $I_G(f)=I_G(g)$

That is, if the norm is unitary-invariant, the whole integration theory will be unitary-invariant.

So that the whole problem of establishing Gaussian integration theory in Hilbert space reduces itself for us to the problem of finding a unitary-invariant norm satisfying the requirements (8.4) (a) and (b). We want to find a norm, moreover, that is as small as possible: for the smaller the norm, the larger the closure $\mathcal{F}$ of the given space $\mathcal{F}_0$ in the norm. How can such a unitary-invariant norm be established?

A first stab would be to write

$$(8.7) \qquad ||f||_0 = \inf_{\substack{g \in \mathcal{F}_0 \\ g(\cdot)\geq|f(\cdot)|}} I_G(g)$$

(by $g(\cdot)\geq|f(\cdot)|$ is meant "$g(x)\geq|f(x)|$" for all x), and to put $|f| = \infty$ if there exists no $g \in \mathcal{F}_0$ such that $g(\cdot)\geq|f(\cdot)|$. This norm evidently is non-negative, while the equations $||af||_0 = |a|\,||f||_0$ and $||f+g||_0 \leq ||f||_0 + ||g||_0$ follow readily by elementary argumentation. Moreover, if $f \in \mathcal{F}_0$ and $g \in \mathcal{F}_0$ satisfies $g(\cdot)\geq|f(\cdot)|$, then $I_G(g)\geq I_G(|f(\cdot)|)\geq|I_G(f)|$ by (8.3) (a),(b) and (c); this shows that $||f||_0 \geq |I_G(f)|$ if $f \in \mathcal{F}_0$. Hence, the norm defined by (8.7), which is evidently unitary-invariant, satisfies the requirements (8.4)(a) and (b), and consequently leads to a unitary-invariant integration theory. This may well be called the theory of "Riemann Integration" over Hilbert space; indeed, the construction defined by a formula exactly like (8.7) and by the

above described process of closure in the corresponding norm would be found on inspection to go from the class of piecewise constant functions defined on finite closed interval of real numbers, and their integrals which may be directly defined as finite sums, to the class of all Riemann-integrable functions defined on this interval, and their Riemann integral.

The deficiency of the above-constructed theory of "Riemann" integration in Hilbert space is, however, even more marked than the well-known deficiency of Riemann integration of functions of a real variable. For instance, take Hilbert space to be represented as the space $\mathcal{L}_2$ of all sequences $\xi=[\xi_1,\xi_2,\ldots]$ such that $\sum_{i=1}^{a} \xi_i^2 < \infty$, and consider the function

$$(8.8) \qquad f(\xi) = \sum_{i=1}^{\infty} c_i \xi_i^2 \text{ , where } c_i>0, \sum_{i=1}^{\infty} c_i < \infty .$$

This quadratic function(a1) has been integrated formally in Appendix III of Chapter I, and the evident integral

$$(8.9) \qquad I_G(f) = \left(\sum_{i=1}^{\infty} c_i \right) \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} y^2 e^{-1/2y^2} dy$$

deduced there. Now, it would be dangerous to deduce the validity of this formula simply from the absolute convergence of the series $\sum_{i=1}^{\infty} c_i$ (Riemann integrals are <u>not</u> countably additive), so that we must investigate the question of whether or not the function defined by (8.8) is actually integrable in the sense that has just been defined. Suppose then that g is a cylinder function such that $g(\xi) \equiv f(\xi)$ for all ξ. Let P be an orthogonal projection with a finite-dimensional range such that $g(P\xi) = g(\xi)$. Let $\xi_0 \neq 0$ be orthogonal to the range of P, so that $P\xi_0 = 0$.

Then

$$g(\xi) = g(P\xi) = g(P(\xi+\lambda\xi_0)) = g(\xi+\lambda\xi_0) ,$$

so that

$$g(\xi) \geq f(\xi+\lambda\xi_0) \to \infty \text{ as } \lambda \to \infty ,$$

so that $g(\xi) = \infty$, which shows that f is not even normable, much less integrable. This shows clearly that the norm defined by (8.7) is too large, and determines an integration theory which is too narrowly applicable. We must consequently search for another, smaller norm.

Suppose that we agree to write $[f_1,f_2,\ldots] > g$ if $f_1,f_2,\ldots$ is a sequence of non-negative integrable cylinder functions such that $\sum_{i=1}^{\infty} f_i(x) \geq g(x)$ for each x. Then the norm

$$\|g\|_1 = \inf_{[f_1,f_2,\ldots]>g} \sum_{i=1}^{\infty} I_G(f_i) \tag{8.10}$$

suggests itself. It is easily seen that this norm satisfies $\|\sum_{i=1}^{\infty} g_i\| \leq \sum_{i=1}^{\infty} \|g_i\|$ for each series $\sum_{i=1}^{\infty} g_i$ of functions, so that the resulting integration theory in Hilbert space has the properties of countable additivity normally associated with the Lebesgue integral. Indeed, since the construction defined by a formula exactly like (8.10) and by the above described process of closure in the corresponding norm would be found on inspection to go from the class of piecewise constant functions defined on an interval of real numbers, and their integrals which may be directly defined as finite sums, to the class of all Lebesgue integrable functions defined on this interval, and their Lebesgue integrals, we should be tempted to call the resulting theory the Gauss-Lebesgue integration theory in Hilbert space. To construct this much more perfect integration theory, we have only to show that $|I_G(f)| \leq \|f\|_1$ for every integrable cylinder function. But, alas! This inequality is absolutely false. Indeed, it was shown in

section 4 of Chapter VI that Hilbert space can be covered in entirety by a set of cylinder sets the sum of whose measures is arbitrarily small. In terms of definition (8.10), this means that $\|f\|_1 = 0$ for every f: every function is a null-function. Clearly then $\|f\|_1$ is <u>too</u> small. Thus the Riemann integration in Hilbert space is too restrictive, the Lebesgue integration is impossible, and we must attempt to find a norm smaller than $\|f\|_0$ but larger than $\|f\|_1$. Such a norm is given by

$$(8.11)\quad \|f\| = \inf_{P\in\Phi,\,\varepsilon>0} \sup_{F\in\Phi_1,\,\|FP-P\|<\varepsilon} (I_G|f(F*(\cdot))|),$$

where Φ_1 is the set of all transformations of $\mathfrak{H}$ into itself having norm at most 1 and finite dimensional range. We must show, first, that this is really a norm and, second, that it has properties (8.4).

Clearly $\|af\| = |a|\,\|f\|$. Let $\|f\|<K_1$ and $\|g\|< K_2$. Then we may find $P_1, P_2 \in \Phi$ and $\varepsilon_1, \varepsilon_2 > 0$ such that

$$(8.12)\quad \sup_{F\in\Phi_1,\,\|FP_1-P_1\|<\varepsilon_1} I_G(|f(F*(\cdot))|) < K_1$$

$$\sup_{F\in\Phi_1,\,\|FP_2-P_2\|<\varepsilon_2} I_G(|g(F*(\cdot))|) < K_2$$

Let P be the projection on the space generated by $P_1\mathfrak{H}$ and $P_2\mathfrak{H}$ and let $\varepsilon = \min\ \varepsilon_1, \varepsilon_2$. Then $PP_1=P_1, PP_2=P_2$. Thus, if $\|FP-P\| < \varepsilon$, we have $\|FP_1-P_1\| = \|(FP-P)P_1\| < \varepsilon\,\|P_1\| = \varepsilon < \varepsilon_1$ and similarly $\|FP_2-P_2\| < \varepsilon_2$. This shows that

$$(8.13)\quad \sup_{F\in\Phi_1,\,\|FP-P\|<\varepsilon} I_G(|f(F*(\cdot))|)< K_1$$

$$\sup_{F\in\Phi_1,\,\|FP-P\|<\varepsilon} I_G(|g(F*(\cdot))|)< K_2$$

so that

$$\sup_{F\in\Phi_1,\ \|FP-P\|<\varepsilon} I_G(|f(F^*(\cdot))+g(F^*(\cdot))|)< K_1 + K_2 ,$$

proving the triangle inequality $\|f+g\|\leq\|f\|+\|g\|$.

All the rest of our argumentation will be based on the following lemma, which is closely related to the "Main Lemma" of Chapter V.

(8.14) Lemma: For each integrable cylinder function f, and each $\delta > 0$, there exists an $\varepsilon > 0$ and a $P \in \Phi$ such that

$$\sup_{F\in\Phi_1,\ \|FP-P\|<\varepsilon} I_G(|f(\cdot)-f(F^*(\cdot))|)< \delta .$$

This lemma will be proved at the end of the present section. Before giving its proof, we shall first show how it may be applied to derive the principal properties of the norm $\|f\|$.

(8.15) Theorem: If f is an integrable cylinder function, then
$\|f\| = I_G(|f(\cdot)|)$.

(8.16) Corollary: If f is an integrable cylinder function, then
$|I_G(f(\cdot))|\leq\|f\|$.

The corollary follows immediately from the theorem by (8.3)(c), and shows that $\|f\|$ has the essential properties (8.4), so that $\|f\|$ may indeed be used to define an extension of the Gauss integral of integrable cylinder functions. To prove the theorem, we argue as follows: let P_0 be an orthogonal projection with finite dimensional range, let Q be the orthogonal projection on the space spanned by $P\mathfrak{H}$ and $P_0\mathfrak{H}$. Then $QP-P = 0$, and $f(Qx) = f(PQx) = f(Px) = f(x)$. Hence $Q \in \Phi_1$ and $\|QP-P\| < \varepsilon$ so that

$$\sup_{F\in\Phi_1,\ \|FP-P\|<\varepsilon} I_G(|f(F^*(\cdot))|)$$

$$\geq I_G(|f(Q(\cdot))|) = I_G(|f(\cdot)|),$$

proving that $\|f\| \geq I_G(|f(\cdot)|)$. On the other hand, if $\delta > 0$, then, by lemma (8.14), we may find $P \varepsilon \underline{\Phi}$ and ε such that

$$\sup_{F \varepsilon \underline{\Phi}_1, \|FP-P\| < \varepsilon} I_G(|f(F^*(\cdot))|) \leq I_G(|f(\cdot)|) + \delta ,$$

so that $\|f\| < I_G(|f(\cdot)|)$ completing the proof of theorem (8.15)

(8.17) Corollary: For any function f we have $\|f\|_0 \geq \|f\|$.

This corollary shows that the integration theory defined as above in terms of the norm $\|f\|$ is indeed more extensive than the above-defined Gauss-Riemann integration theory in Hilbert space. To prove it, note that if g is a cylinder function such $g(\cdot) \geq |f(\cdot)|$, then it follows immediately from definition (8.11) that $\|g\| \geq \|f\|$. On the other hand, by Theorem (3.15), $\|g\| = I_G(g(\cdot))$. Thus $\|f\|_0 = \inf I_G(g(\cdot)) \geq \|f\|$, proving the corollary.

Another theorem of interest may be stated as follows:

(8.18) Theorem: If f is integrable, and $\delta > 0$, there exists an $\varepsilon > 0$ and a $P \varepsilon \underline{\Phi}$ such that

$$\sup_{F \varepsilon \underline{\Phi}_1, \|FP-P\| < \varepsilon} \|f(\cdot) - f(F^*(\cdot))\| < \delta.$$

(8.19) Corollary: Let f be integrable, and let P_n be a sequence of projections such that $P_n x \longrightarrow x$ as $n \longrightarrow \infty$ for each $x \varepsilon h$. Then, for all sufficiently large n, $f(P_n(\cdot))$ is an integrable cylinder function, and

$$\text{(i)} \quad \lim_{\min n \longrightarrow \infty} I_G(|f(P_n(\cdot)) - f(P_m(\cdot))|) = 0$$

$$\text{(ii)} \quad \lim_{n \longrightarrow \infty} I_G(f(P_n(\cdot))) = I_G(f)$$

The corollary shows that every function integrable in the integration theory defined in terms of the norm $\|f\|$ as above is "completely invariant" in the sense of s ction 5 of Chapter V. To prove it from the theorem, we have only to show that $\|P_n P - P\| = \sup_{\|x\| \leq 1, x \varepsilon \mathfrak{H}} \|P_n Px - Px\| < \varepsilon$ for each $P \varepsilon \underline{\Phi}$. This evidently amounts to showing that $\sup_{\|x\| \leq 1, x \varepsilon P\mathfrak{H}} \|P_n x - x\| < \varepsilon$

for sufficiently large n. Since we have assumed that $\lim_{n\to\infty} \| P_n x - x \| = 0$ for each x, and P_1 is finite-dimensional, this last equation follows readily.

To prove Theorem (8.18), we argue as follows. Since f is integrable, there exists a cylinder function g such that $\| f-g \| < \delta/10$. Thus, there exists an $\varepsilon_1 > 0$ and $P_1 \in \Phi$ such that

$$(8.20) \quad \sup_{F\in\Phi_1,\ \| FP_1-P_1 \| < \varepsilon_1} I_G(|f(F(\cdot))-g(F(\cdot))|) < \delta/10$$

Let $\tilde{F} \in \Phi_1$, and suppose that $\| \tilde{F}P_1-P_1 \| < \varepsilon_2 = 1/2\varepsilon_1$, $F \in \Phi_1$ and $\| FP_1-P_1 \| < \varepsilon_2 = 1/2\varepsilon_1$. Then $\| F\tilde{F}P_1-P_1 \| \le \| F\tilde{F}P_1-\tilde{F}P_1 \| + \| \tilde{F}P_1-P_1 \| \le \varepsilon_2+\varepsilon_2 = \varepsilon_1$. Hence, by (8.20), we have

$$(8.21) \quad \sup_{F\in\Phi_1,\ \| FP_1-P_1 \| < \varepsilon_2} I_G(|f(\tilde{F}*F*(\cdot))-g(\tilde{F}*F*(\cdot))|) < \delta/10 .$$

Using Lemma (8.14), we find that there exists a projection $P_2 \in \Phi$ and an $\varepsilon_3 > 0$ such that

$$(8.22) \quad \sup_{F\in\Phi_1,\ \| FP_2-P_2 \| < \varepsilon_3} I_G(|g(\cdot) - g(F*(\cdot))|) < \delta/10 .$$

In equations (8.20), we may replace P_1 by any projection P such that $PP_1 = P_1$, and replace ε_1 by any smaller positive number. Suppose, indeed, that we let P be the projection onto the space spanned $P_1\mathcal{H}$ and $P_2\mathcal{H}$ and let $\varepsilon = 1/2 \min(\varepsilon_1,\varepsilon_3)$. Then $\| \tilde{F}P-P \| < \varepsilon$ and $\| FP-P \| < \varepsilon$ together imply $\le \| \tilde{F}(FP-P) \| + \| \tilde{F}P-P \| < 2\varepsilon \le \varepsilon_3$, so that by (8.20), (8.21) and (8.22) we have

(8.23)
$$\sup_{F\epsilon\Phi_1,\ \|FP-P\|<\epsilon} I_G((|f(F^*(\cdot))-g(F^*(\cdot))|)<\delta/10$$

$$\sup_{F\epsilon\Phi_1,\ \|FP-P\|<\epsilon} I_G(|f(\tilde{F}^*F^*(\cdot))-g(\tilde{F}^*F^*(\cdot))|)<\delta/10$$

$$\sup_{F\epsilon\Phi_1,\ \|FP-P\|<\epsilon} I_G(|g(\tilde{F}^*F^*(\cdot))-g(\cdot)|)<\delta/10$$

$$\sup_{F\epsilon\Phi_1,\ \|FP-P\|<\epsilon} I_G(|g(F^*(\cdot))-g(\cdot)|)<\delta/10$$

provided that $\tilde{F}\epsilon\Phi_1$ and $\|\tilde{F}P-P\|<\epsilon$. Adding these four inequalities, we find that

$$\sup_{F\epsilon\Phi_1,\ \|FP-P\|} I_G(|f(\tilde{F}^*F^*(\cdot))-f(F^*(\cdot))|)<\delta$$

if $F\epsilon\Phi_1$ and $\|FP-P\|<\epsilon$, i.e., that

$$\|f(\tilde{F}(\cdot))-f(\cdot)\|<\delta$$

if $\tilde{F}\epsilon\Phi_1$ and $|\tilde{F}P-P|<\epsilon$. This proves Theorem(8.18).

At this point, let us list various useful properties of I_G, of which (a) and (d) have already been proved.

(8.24) (a) The set of all integrable functions is a linear space, and the Gaussian integral I_G is a linear functional defined on this space.

(b) $I_G(1) = 1$

(c) If $f \geq 0$ and f is integrable then $I_G(f) \geq 0$.

(d) If f is integrable, and $g(x) = f(U(x)$ where U is a unitary mapping of h onto itself, then g is integrable and $I_G(g) = I_G(f)$.

(e) If f is integrable, and ϕ is a function of a complex variable satisfying a Lipschitz condition in the whole plane, then ϕ (f) is integrable.

(f) If f is integrable, so is $|f(\cdot)|$, and $I_G(|f(\cdot)|) \geq |I_G(f)|$.

(h) If f is integrable and bounded, and g is integrable, then fg is integrable.

To prove those of these statements which have not already been proved, we may argue as follows.

Since the function identically equal to 1 is a cylinder function, (b) follows immediately from definition (8.1) of the integral of a cylinder function. Let f satisfy the hypotheses of (c), and let g_n be a sequence of integrable cylinder functions such that $\| g_n - f \| \longrightarrow 0$ as $n \longrightarrow \infty$. Let h_n be defined by $h_n(x) = |g_n(x)|$. Then $|f(x)-h_n(x)| \leq |f(x)|$ for every x. Thus it follows from definition (8.11) that $\| f-h_n \| \leq \| f-g_n \|$, so that $\| f-h_n \| \longrightarrow 0$. Since h_n is an integrable cylinder function for each n we have $I_G(f) = \lim_{n\longrightarrow\infty} I_G(h_n)$ by definition. Thus, since $I_G(h_n) \geq 0$ by (8.3)(a), $I_G(f) \geq 0$ proving(c).

Next, let $|\phi(z_1)-\phi(z_2)| \leq K|z_1-z_2|$ for each pair z_1, z_2 of complex numbers as in (e), let f satisfy the hypotheses of (e), and let g_n be a sequence of integrable cylinder functions such that $\| g_n - f \| \longrightarrow 0$ as $n \longrightarrow \infty$. The functions h_n defined by $h_n(x) = \phi\ (f_n(x))$ are integrable cylinder functions, and $|\phi(f(x))-\phi(g_n(x))| \leq K|f(x)-g_n(x)|$ for each x. Thus it follows from definition (8.11) that $\| \phi f - h_n \| \leq K \| f-g_n \|$, so that $\| \phi\ f-h_n \| \longrightarrow 0$ as $n \longrightarrow \infty$, proving that ϕf is integrable, and establishing (e).

The integrability of $|f(\cdot)|$ under the hypotheses of (f) follows from (e), and the fact that $I_G(|f(\cdot)|)$ is larger than the real part of $I_G(af) = aI_G(f)$ for each a of modulus 1 follows

from (c) and the fact that $|f(x)|$ is larger than the real part of $af(x)$ for each x and each a of modulus 1. This establishes (f).

Finally, to prove (h), suppose that f is a bounded integrable function, and that $|f(x)| \leq K < \infty$ for all x. Let g be integrable; let $\{f_n\}$ be a sequence of integrable cylinder functions such that $\|f_n - f\| \longrightarrow 0$, and let $\{g_m\}$ be a sequence of integrable cylinder functions such that $|g_m - g| \longrightarrow 0$. Using a well-known theorem about finite-dimensional integrals, we may find, for each m, a bounded integrable cylinder function $\tilde{g}_m$ such that $I_G(|g_m(\cdot) - \tilde{g}_m(\cdot)|) \leq 1/m$. It then follows immediately from Lemma (8.14) and definition (8.11) that $\| g_m - \tilde{g}_m \| \leq 1/m$. Hence $\| \tilde{g}_m - g \| \longrightarrow 0$. Let K_m be the least upper bound of $|\tilde{g}_m(x)|$ as x ranges over h. Then

$$|f(x)g(x) - f_n(x)\tilde{g}_m(x)| \leq |f(x) - f_n(x)|\,|\tilde{g}_m(x)| + |f(x)|\,|g(x) - \tilde{g}_m(x)|$$

$$\leq K_m |f(x) - f_n(x)| + K|g(x) - \tilde{g}_m(x)|$$

Thus, by definition (8.11), $\|f(\cdot)g(\cdot) - f_n(\cdot)\tilde{g}_m(\cdot)\|$

$$\leq K_m \|f - f_n\| + K\|g - \tilde{g}_m\|$$

Let $\varepsilon > 0$; choose m such that $K\| g - \tilde{g}_m \| < \varepsilon/2$, and then choose n such that $K_m \|f - f_n\| < \varepsilon/2$. Then $\|f(\cdot)g(\cdot) - f_n(\cdot)\tilde{g}_m(\cdot)\| < \varepsilon$. Since $f_n(\cdot)\tilde{g}_m(\cdot)$ is evidently an integrable cylinder function, it follows that $f(\cdot)g(\cdot)$ is integrable. Thus all the statements of (8.24) are proved.

To show that the integration theory defined in terms of the norm (8.11) is in fact significantly stronger than the Riemann-Gauss integration theory defined in terms of the norm (8.7), let us re-examine the example given in (8.8) above, which may be written in a unitary-invariant fashion as

(8.25) $\quad f(x) = \sum_{i=1}^{\infty} c_i (x, v_i)^2, c_i > 0, \sum_{i=1}^{\infty} c_i < \infty$, where v_i is

an orthonormal basis for Hilbert space. We saw above that $\|f\|_0 = \infty$. On the other hand

$$\|f\|_1 \leq \sup_{F\epsilon\Psi_1} \int_{\mathfrak{H}} \left\{ \sum_{i=1}^{\infty} c_i (Fx, v_i)^2 \right\} \mu_G(dx)$$

$$\leq \sup_{F\epsilon\Psi_1} \sum_{i=1}^{\infty} c_i \int_{\mathfrak{H}} (x, F^* v_i)^2 \mu_G(dx)$$

$$= (\sqrt{2\pi})^{-1} \int_{-\infty}^{+\infty} x^2 \, e^{-1/2\, x^2} \, dx \sup_{F\epsilon\Psi_1} \sum_{i=1}^{\infty} c_i |F^* v_i|^2$$

$$\leq \left(\sum_{i=1}^{\infty} c_i \right) \left(\sqrt{2\pi} \int_{-\infty}^{+\infty} x^2 e^{-1/2\, x^2} dx \right) < \infty \quad ;$$

here the interchange of integration and summation between the first and second lines is permissible since finite dimensional integrals are countably additive, i.e., since

$$\sum_{i=1}^{\infty} \int_{\mathfrak{H}} g_i(Fx) \mu_G(dx) = \int_{\mathfrak{H}} \sum_{i=1}^{\infty} g_i(Fx) \mu_G(dx)$$

for each series of non-negative integrable functions g by definition (8.1) and by the Lebesgue theorem for finite dimensional integrals. In the same way we may show that

$$\| f(\cdot) - \sum_{i=1}^{n} c_i \, (\cdot v_i)^2 \| = 0 \, \left(\sum_{i=n+1}^{\infty} c_i \right) \longrightarrow 0 \, ,$$

so that f is indeed integrable, and in fact

$$I_G(f) = (2\pi)^{-1/2} \left(\lim_{n\to\infty} \sum_{i=1}^{n} c_i \right) \int_{-\infty}^{+\infty} x^2 e^{-1/2\, x^2} dx = \sum_{i=1}^{\infty} c_i$$

as was to be hoped.

The validity of a number of formal evaluations of Gauss integrals in Hilbert space may be established in the same way. One interesting formula is as follows: let A and B be compact Hermitian operators. Suppose that the sum of the absolute values of the eigenvalues (counted with proper multiplicities) of both A and B converge, and that $B < 1/2\ I$. Then

$$(8.26)\quad \int_{\mathfrak{H}} (Ax,x)\exp\left\{(Bx,x)\right\}\mu_G(dx) = \mathrm{tr}(A(I-2B)^{-1})\det((I-2B)^{-1/2})$$

in terms of the trace tr and determinant det of a pair of infinite matrices. We shall give a suggestive formal derivation of (8.26) here, leaving the justification of the following formal computations to the reader as an exercise.

$$\int_{\mathfrak{H}} (Ax,x)\exp\left\{(Bx,x)\right\}\mu_G(dx) = (2\pi)^{-\frac{\infty}{2}}\int_{\mathfrak{H}} (Ax,x)\exp\left\{-\tfrac{1}{2}(x,x)+(Bx,x)\right\}dx$$

$$= (2\pi)^{-\frac{\infty}{2}}\int_{\mathfrak{H}} (Ax,x)\exp\left\{-\tfrac{1}{2}(I-2B)x,x)\right\}dx$$

$$= (2\pi)^{-\frac{\infty}{2}}\int_{\mathfrak{H}} (Ax,x)\exp -\tfrac{1}{2}\|(I-2B)^{\frac{1}{2}}x\|^2\,dx$$

$$= (2\pi)^{-\frac{\infty}{2}}\det\left\{I-2B)^{-\frac{1}{2}}\right\}\int_{\mathfrak{H}_y} (A(I-2B)^{-\frac{1}{2}}y,(I-2B)^{-\frac{1}{2}}) \cdot e^{-\frac{1}{2}\|y\|^2}\,dy,$$

making the change of variable $y = (I-2B)^{1/2}x$ and using the ordinary determmnant formula for change of variables in a multiple integral. Thus,

$$\int_{\mathfrak{H}} (Ax,x)\exp\left\{(Bx,x)\right\}\mu_G(dx) = \det[(I-2B)^{-\frac{1}{2}}]\int_{\mathfrak{H}} (I-2B)^{-\frac{1}{2}}A(I-2B)^{-\frac{1}{2}}y,y)\cdot \mu_G(dy)\cdot$$

If the Hermitian operator $(I-2B)^{-\frac{1}{2}}A(I-2B)^{-\frac{1}{2}}$ is reduced to diagonal form by a rotation of Hilbert space, this last integral can be written as the sum of two integrals of the form (8.25). Hence, making use of the value found for an integral of the form (8.25) we obtain

$$\int_{\mathfrak{H}} (Ax,x)\exp\left\{(Bx,x)\right\}\mu_G(dx) = \mathrm{tr}((I-2B)^{-\frac{1}{2}}A(I-2B)^{-\frac{1}{2}})\det\left\{(I-2B)^{-\frac{1}{2}}\right\}$$

Since $\mathrm{tr}(CD) = \mathrm{tr}(DC)$, (8.26) follows immediately.

We conclude the present section by giving the proof of the fundamental lemma (8.14).

Let f be an integrable cylinder function, so that $f(x) = \phi((x,v_1),\ldots,(x,v_n))$, where $\{v_1,\ldots,v_n\}$ is an orthonormal set of n vectors, and where ϕ is a function of n variables such that

$$(2\pi)^{-\frac{n}{2}} \int_{E^n} \phi(x_1,\ldots,x_n) e^{-\frac{1}{2}(x_1^2+\ldots+x_n^2)} dx_1\ldots dx_n$$

exists. Let F be the projection on the space spanned by $v_1,\ldots,v_n$, let S be the 2n dimensional space spanned by $v_1,\ldots,v_n$ and n other, linearly independent, vectors, and let $F_\varepsilon \varepsilon \mathfrak{F}_{-1}$, $|F_\varepsilon P - P| < \varepsilon$. We must show that

$$(8.27)\quad \lim_{\varepsilon \to 0} \int_{\mathfrak{H}} |f(F_\varepsilon^*(x)) - f(x)| M_G(dx) = 0 ,$$

i.e., that

$$(8.28)\quad \lim_{\varepsilon \to 0} \int_{\mathfrak{H}} |\phi((x,F_\varepsilon v_1),\ldots,(x,F_\varepsilon v_n)) - \phi((x,v_1),\ldots,(x,v_n))| \mu_G(dx) = 0$$

For each ε, let U_ε be a unitary mapping of Hilbert space onto itself such that $U_\varepsilon v_i = v_i, i=1,\ldots,n$, and such that $U_\varepsilon F_\varepsilon v_i \varepsilon S$, $i = 1,\ldots,n$. Let $U_\varepsilon F_\varepsilon = L_\varepsilon$; then since $U_\varepsilon v_i = v_i$, $i=1,\ldots,n$, we have $U_\varepsilon P = P$, so that $|L_\varepsilon P - P| = |U_\varepsilon F_\varepsilon P - U_\varepsilon P| = |F_\varepsilon P - P| < \varepsilon$. Moreover, by the unitary invariance of the Gauss integral for cylinder functions, (8.28) is equivalent to

$$(8.29)\quad \lim_{\varepsilon \to 0} \int_{\mathfrak{H}} |\phi((x,L_\varepsilon v_1),\ldots(x,L_\varepsilon v_n)) - \phi((x,v_1),\ldots,(x,v_n))| \mu_G(dx)$$

If Q denotes the orthogonal projection on the 2n dimensional space s, then

$$\phi((Qx,L_\varepsilon v_1),\ldots,(Qx,L_\varepsilon v_n)) = \phi((x,L_\varepsilon v_1),\ldots,(x,L_\varepsilon v_n))$$

and

$$\phi((Qx,v_1),\ldots,(Qx,v_n)) = \phi((x_1,v_1),\ldots,(x_n,v_n)) .$$

Thus by definition (8.1), we may write (8.29) in equivalent form as

(8.30)
$$\overline{\lim_{\varepsilon \to 0}} \int_S |\phi((x,L_\varepsilon v_1),\ldots(x,L_\varepsilon v_n))-\phi((x,v_1),\ldots(x,v_n)|e^{-\frac{1}{2}\|x\|^2}dx < \delta$$

For each $\delta > 0$ let ψ be a bounded continuous function of n real variables such that

(8.31)
$$\int_{E^n} |\psi(x_1,\ldots,x_n)-\phi(x_1,\ldots,x_n)|e^{-\frac{1}{2}(x_1^2+\ldots+x_n^2)}dx_1\ldots dx_n < \frac{\delta}{3}$$

Then, since
$$\overline{\lim_{\varepsilon \to 0}} \int_S |\psi((x,L_\varepsilon v_1),\ldots,(x_1L_\varepsilon x_n)-\psi((x,v_1),\ldots(x,v_n))|e^{-\frac{1}{2}\|x\|^2}dx=0$$

by the boundedness and continuity of ψ and by the Lebesgue dominated convergence theorem, it follows immediately from(8.31) that (8.30) is a consequence of

(8.32)
$$\overline{\lim_{\varepsilon \to 0}} \int_S |\phi((x,L_\varepsilon v_1),\ldots,(x,L_\varepsilon v_n))-\psi((x,L_\varepsilon v_1),\ldots,(x,L_\varepsilon v_n))|e^{-\frac{1}{2}\|x\|^2}dx < \frac{\delta}{3}$$

But (8.32) will be an easy consequence of (8.31), once the following auxillary lemma is proved.

(8.33) Lemma: If h is a Gauss-integrable function on a finite-dimensional Hilbert space S, and M is a 1-1 linear mapping of S into itself such that $\|M\| \leq 1$, then
$$\int_S |h(Mx)|e^{-\frac{1}{2}\|x\|^2}dx \leq |\det(M)|^{-1}\int_S |h(x)|e^{-\frac{1}{2}\|x\|^2}dx$$

Proof of lemma(8.33): Make the change of variables $Mx = y$. Then $\|y\|^2 \leq \|x\|^2$, so that $e^{-\frac{1}{2}\|y\|^2} \geq e^{-\frac{1}{2}\|x\|^2}$. Hence

$$\int_S |h(Mx)|e^{-\frac{1}{2}\|x\|^2}dx \le \int_S |h(y)|e^{-\frac{1}{2}\|y\|^2}dx$$

$$= |\det(M)|^{-1}\int_S |h(y)|e^{-\frac{1}{2}\|y\|^2}dy ,$$

Q. E. D.

Proof of (8.32) from Lemma(8.33): Find any convenient vectors $u_{n+1},\dots,u_{2n}$ so that $\{v_1,\dots,v_{2n}\}$ is an orthonormal basis for S. Since $\|L_\varepsilon v_i - v_i\| \to 0$ as $\varepsilon \to 0$ for $i=1,\dots n$, the set $K_\varepsilon = \{L_\varepsilon v_1,\dots,L_\varepsilon v_n, v_{n+1},\dots,v_{2n}\}$ is linearly independent for all sufficiently small ε. Using the Gram - Schmidt orthonormalisation process, construct an orthonormal basis $\{u_1^{(\varepsilon)},\dots,u_{2n}^{(\varepsilon)}\}$ for S out of the set K_ε. Now put $M_\varepsilon(\sum_{i=1}^{2n} a_i v_i) = \sum_{i=1}^{n} a_i L_\varepsilon v_i + \sum_{i=n+1}^{2n} a_i u_i^{(\varepsilon)}$. Then

$$\|M_\varepsilon(\sum_{i=1}^{2n} a_i v_i)\|^2 = \|L_\varepsilon(\sum_{i=1}^{n} a_i v_i)\|^2 + \sum_{i=n+1}^{2n} a_i^2$$

$$\le \|\sum_{i=1}^{n} a_i v_i\|^2 + \sum_{i=n+1}^{2n} a_i^2 = \sum_{i=1}^{2n} a_i^2$$

Hence $\|M_\varepsilon\| \le 1$. Moreover, for ε sufficiently small, M_ε is evidently 1-1. Putting

$$h(x) = \phi((x,v_1),\dots,(x,v_n)) - \psi((x,v_1),\dots,(x,v_n))$$

for $x \in S$, (8.31) states

$$(8.34)\quad \int_S |h(x)|e^{-\frac{1}{2}\|x\|^2}dx < \frac{\delta}{3}$$

and (8.32) states

$$(8.35)\quad \overline{\lim_{\varepsilon\to 0}} \int_S |h(M_\varepsilon^* x)|e^{-\frac{1}{2}\|x\|^2}dx < \frac{\delta}{3}$$

But(8.35) follows from (8.34) by Lemma(8.33). Thus the fundamental lemma(8.14) is completely proved.

Q. E. D.

Chapter IX

The Wiener Integral

§ 1. Introduction

We now wish to give a brief exposition of the ideas related to the notion of Wiener's integral. Wiener's approach to the problem of integrating over a function space consists, in essence, of establishing a totally additive measure for the function space via the principle of transference. That is, a finitely additive measure is assigned to certain selected "intervals" in the function space and these "intervals" are mapped onto certain subintervals of the closed interval [0,1] in such a way that the length of the image interval (its Lebesgue measure) equals the assigned measure of the original interval. Thus correspondence of intervals is used to induce an "almost" 1 - 1 map of the function space and the interval [0,1]. This mapping, in turn, induces the desired measure theory in the function space - a set of functions being measurable if the image set is a Lebesgue measurable subset of [0,1] - and the measure of the set is defined to be the Lebesgue measure of the image set. The measure on the function space obtained in this way is then totally additive. The mapping between the function space and [0,1] also induces a mapping of functionals over the function space into functions over [0,1] in such a way that one may then define the integral of a functional as the Lebesgue integral of its image function over [0,1].

§ 2. Wiener Measure.

The space of functions which we consider initially is the space W of all functions $x(\cdot)$ defined on [0,1] and having $x(0) = 0$.

Definition: Let $0 < t_1 < t_2 < \ldots < t_n \leq 1$ be an arbitrary subdivision of the unit interval, and let $\{a_i, b_i\}$ be arbitrary real numbers with $a_i < b_i$, $(i=1,\ldots n)$. The subset of W

defined by the conditions

$$a_i \leq x(t_i) \leq b_i \qquad (i = 1,\ldots,n)$$

is called a (closed) quasi-interval.

To avoid confusion with the Lebesgue measure on the real line, which will be used later, the measure which we shall assign to a quasi-interval will be called the probability of the quasi-interval I (i.e. the probability that $x(\cdot) \in I$), and shall have the value

$$P(I) = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} \prod_{\nu=1}^{n} \left\{ \exp\left[-\frac{1}{2} \frac{(\xi_\nu - \xi_{\nu-1})^2}{\Delta_\nu} \right] \frac{d\xi_\nu}{\sqrt{2\pi\Delta_\nu}} \right\}$$

where $\xi_o = t_o = 0$ and $\Delta_\nu = t_\nu - t_{\nu-1}$. From the usual "compatibility" conditions satisfied by the Gaussian measures it follows that this assigned probability is finitely additive with respect to the quasi-intervals.

As was hinted earlier one does not attempt, initially, to determine a probability function over the function space which assigns a probability to every quasi-interval. Rather, one constructs a function which will assign the given measure to a certain set of quasi-intervals so that one can define the final probability function by "continuity".

To this end we consider a sequence of partitions. $\{\rho^{(n)}\}$ of the real line such that

(1) each partition divides the real line into a finite number of intervals (two of them unbounded),

(2) for each n, $\rho^{(n+1)}$ is a refinement of $\rho^{(n)}$, and

(3) as $n \to \infty$ the set of division points $\{K_\nu^{(n)}\}$ becomes dense on the whole real line. For example, let $\rho^{(n)}$ consist of the $(2n^2+2)$ intervals

$$(-\infty, -n], [-n, -n+\tfrac{1}{n}], \ldots, [-\tfrac{1}{n}, 0], [0, \tfrac{1}{n}], \ldots, [n-\tfrac{1}{n}, n], [n, \infty) .$$

We also consider a monotonically increasing sequence of finite

sets $\left\{t_j^{(n)}\right\}$ becoming dense in [0,1]. For example, let $t_j^{(n)} = j2^{-n}$. Finally we consider the set of all quasi-intervals of the form

$$\left\{x(\cdot):\ x(t_j^{(n)}) \in A_j^{(n)} \quad \text{for each } j\right\}$$

where each $A_j^{(n)}$ is one of the intervals of $\rho^{(n)}$. Note that, for a given n, these quasi-intervals are mutually disjoint and together exhaust W. Note further that, if $\bar{n} > n$, each of the quasi-intervals associated with n is the union of a subset of the quasi-intervals associated with $\bar{n}$. We denote the quasi-intervals for a given n by $\left\{I_j^{(n)}\right\}$, ordering them recursively so that those of the quasi-intervals which cover $I_1^{(n-1)}$ come first followed by those which cover $I_2^{(n-1)}$ and so on.

We are now ready to construct our mapping T of W into [0,1]. Let

$$\tau_j^{(n)} = \sum_{\nu \le j} P(I_\nu^{(n)}) \quad \text{and} \quad \mathcal{J}_j^{(n)} = [\tau_{j-1}^{(n)}, \tau_j^{(n)}] \ .$$

Then $\mathcal{J}_j^{(n)}$ shall be the image under our map of $I_j^{(n)}$. We must now show, as was mentioned in Section §1, that T is "almost" 1-1. To be more precise, we shall show that - except for the denumerable set of endpoints of intervals $\mathcal{J}_\nu^{(n)}$ - the pre-image under T of any $\alpha \in [0,1]$ is a set of functions uniquely determined by their values on the dense set $\{t_j\} = \lim_{n\to\infty} \left\{t_\nu^{(n)}\right\}$. If we then restrict ourselves to the set C of those functions in W which are continuous, T will thus be 1-1 almost everywhere. It will be clear that the map T, considered as mapping C into [0,1], is, almost everywhere, uniquely defined independently of the sequences $\left\{\rho^{(n)}\right\}$ and $\left\{t_j^{(n)}\right\}$ employed in the construction.

Let α be any point of [0,1] which is not the endpoint of an $\mathcal{J}_\nu^{(n)}$. Then, as the intervals of the $\rho^{(n)}$ are becoming dense, the measure of each of the $I_j^{(n)}$ approaches zero with increasing n from which there follows the existence of a unique nested sequence $\left\{\mathcal{J}_{\nu_n}^{(n)}\right\}$ such that $\alpha \in \mathcal{J}_{\nu_n}^{(n)}$ for n = 1,2,... This sequence

corresponds to the sequence $\left\{I_{\nu_n}^{(n)}\right\}$ which, for each t_j in the dense set $\{t_j\}$, determines a nested sequence of $\rho^{(n)}$-intervals. By condition (3) on $\rho^{(n)}$ this sequence uniquely determines $x(t_j)$. We have, therefore, shown that the pre-image under T of each a contains <u>at most</u> one function $x(\cdot) \in C$.

It is not true that the pre-image of every a contains <u>at least</u> one continuous function. We shall show, however, that the set of points in the interval [0,1] which have no pre-image in C has Lebesgue measure zero. Thus we will have a 1-1 measure preserving (for the quasi-intervals) map of part of C onto a subset of [0,1] having measure one, thus completing the transference.

<u>Definition</u>: A set $S \subset W$ is called a <u>null set</u> if, for any given $\varepsilon > 0$, there exists a countable collection of quasi-intervals which cover S, the sum of whose probabilities is less than ε. Clearly the set of all $a \in [0,1]$ which have pre-images in a given null set S has Lebesgue measure 0.

The statement that the preimage in W of a does not contain a continuous function is equivalent to asserting that the $\{x(t_j)\}$ determined by a (assuming that a is not the end point of an image interval) do not constitute a continuous function "relative to $\{t_j\}$". Hence, since the set of endpoints of image intervals is countable (and, therefore, constitute a set of measure 0) it will more than suffice to show that the $x(t_j)$ which are not Hölder continuous $(h, \frac{1}{4})$ for any h constitute a null set. Thus we need only prove the

<u>Lemma</u>: The set $S \subset W$ consisting of those functions such that for every $h > 0$ there exists a pair of binary rationals t_1 and t_2 (depending on h and on the function) such that

$$|x(t_1) - x(t_2)| > \left(\frac{2h}{2^{1/4} - 1}\right)|t_1 - t_2|^{1/4} \tag{2.1}$$

is a null set.

Proof: Denoting by S_h the set of all functions such that (2.1) is satisfied for this fixed h and some t_1 and t_2, we observe immediately that $S_{h_1} \supset S_{h_2}$ if $h_2 > h_1$. Thus $S = \bigcap_{h>0} S_h$ and $S \subset S_h$ for every $h > 0$, and it suffices, therefore, to show that for any given $\varepsilon > 0$ there exists an h such that S_h is contained in a set of quasi-intervals the sum of whose probabilities is not greater than ε.

We wish to show, first, that if x satisfies (2.1) there exists a pair of integers m,k such that

$$(2.2) \qquad \left| x\left(\frac{m}{2^k}\right) - x\left(\frac{m+1}{2^k}\right) \right| > h\left(\frac{1}{2^k}\right)^{1/4} .$$

Assume $t_2 > t_1$. Since t_1 and t_2 are binary rationals we may write

$$\frac{1}{2}(t_2 - t_1) = \sum_{k=1}^{\infty} \frac{c_k}{2^k} ,$$

where each $c_k = 0$ or 1, and all but a finite number of $c_k = 0$. Let c_j be the first, and C_K the last, non-zero coefficient and set $t_3 = \frac{1}{2}(t_1+t_2)$. Then

$$(2.3) \qquad |x(t_1) - x(t_2)|$$
$$\leq |x(t_1)-x(t_1+2^{-j})| + \ldots + \left|x\left(t_1 + \sum_{k=j}^{K-1} \frac{c_k}{2^k}\right) - x(t_3)\right|$$
$$+ |x(t_3)-x(t_3+2^{-j})| + \ldots + \left|x\left(t_3 + \sum_{k=j}^{K-1} \frac{c_k}{2^k}\right) - x(t_2)\right| .$$

Each of the terms on the right side of the inequality (2.3) is of the form of the left side of (2.2). If (2.2) is false for <u>all</u> m,k we would have

$$|x(t_1)-x(t_2)| \leq h \cdot 2 \sum_{k=j}^{K} 2^{-k/4} < \frac{2h\, 2^{-j/4}}{1 - 2^{-1/4}}$$

whereas

$$|t_1-t_2|^{1/4} = \left| 2 \sum_{k=j}^{K} \frac{c_k}{2^k} \right|^{1/4} \geq 2^{1/4}\, 2^{-j/4}$$

which yields a contradiction of (2.1).

Thus it suffices to show that, for any $\varepsilon > 0$ and for h sufficiently large, the set of functions which satisfy (2.2) for some m,k can be included in the union of a countable collection of quasi-intervals the sum of whose probabilities is less than ε. Let m and k be fixed. Since our selected quasi-intervals divide the range of values of $x(m2^{-k})$ and $x((m+1)2^{-k})$ to an arbitrary degree of fineness we can cover the set of functions satisfying (2.2) by a finite collection of quasi-intervals the sum of whose probabilities is less than

$$\frac{2}{\sqrt{\pi 2^{-k}}} \int_{h2^{-k/4}}^{\infty} \exp[-x^2/2^{-k}]\, dx + \frac{\varepsilon}{2^{2k+2}}$$

$$= \frac{\varepsilon}{2^{2k+2}} + \frac{2}{\sqrt{\pi}} \int_{h2^{k/4}}^{\infty} e^{-x^2}\, dx$$

$$< \frac{\varepsilon}{2^{2k+2}} + \frac{2}{\sqrt{\pi}} \exp[-h2^{k/4}] \quad .$$

If we sum over m and k we see that the sum of the probabilities of the countable number of covering quasi-intervals is less than

$$\frac{\varepsilon}{2} + \frac{4}{\sqrt{\pi}} \sum_{k=1}^{\infty} 2^k \exp[-h2^{k/4}]$$

which is, in turn, less than ε for h sufficiently large. Q.e.d.

§ 3. The Wiener Integral

The measure defined above on the space of continuous functions $x(\cdot)$ on [0,1] such that $x(0) = 0$ is called the Wiener integral. Because of the way in which this measure was brought into existence in the previous section integration over C with respect to this measure can be given a particularly simple interpretation in terms of Lebesgue integration over the interval [0,1].

The construction of the previous section provides a measure preserving map between C and [0,1] which is 1-1 apart from sets of measure 0 in each. Under this map let

$$x(\cdot) \longleftrightarrow a \quad .$$

Then if F[x] is a functional over C there corresponds a function

$$\phi(a) = F[T^{-1}(a)] = F[x(\cdot)]$$

which is defined almost everywhere on [0,1]. F is called Wiener integrable if the corresponding ϕ is Lebesgue integrable and the Wiener integral is then given by

$$\int_C F[x]\, d_W x = \int_0^1 \phi(a)da \quad .$$

Chapter X: Remarks about the Wiener Measure

It was shown earlier (in Chapter IV) that the integral over the Hilbert space, as we had introduced it in Chapter I , does not lead to a totally additive measure. If one wants the measure to have this property one must extend the class of independent functions beyond the Hilbert space. Such extensions can easily be given.

For example, suppose the independent function is a sequence $\phi = \{\phi_1, \phi_2, \ldots\}$ with the unit form

$$(\phi,\phi) = \phi_1^2 + \phi_2^2 + \ldots \quad .$$

Then the measure with the differential

$$e^{-\frac{1}{2}(\phi,\phi)} \prod_\nu \frac{d\phi_\nu}{\sqrt{2\pi}}$$

can be made totally additive by admitting all sequences as independent functions without imposing the restriction $\sum_\nu \phi_\nu^2 < \infty$. It is not necessary, though, to employ such a radical extension and the question arises: which less radical extensions are sufficient to make the measure totally additive.

Such a question will be discussed in Chapter XIV in connection with the measure differential which may be given symbolically as

$$e^{-\frac{1}{2}(\phi,\phi)} \prod_s \sqrt{\frac{ds}{2\pi}}\, \delta\phi(s) \ ,$$

where $\phi(\cdot)$ is a function of s for $0 \le s \le 1$ and

$$(\phi,\phi) = \int_0^1 \phi^2(s)ds.$$

The Wiener measure can be obtained from the measure so introduced by expressing it in terms of the function χ defined by

$$\chi(s) = \int_0^s \phi(s')ds' \quad .$$

Accordingly, the Wiener measure may, symbolically, be written as

$$\exp\left[-\frac{1}{2}\int_0^1 \frac{(d\chi(s))^2}{ds}\right]\prod_{s\neq 0}\frac{\delta\chi(s)}{\sqrt{2\pi ds}} \quad .$$

The Wiener integral of a function $k[\chi]$ can then be written symbolically as

$$\int^{\infty} k[\chi]\exp\left[-\frac{1}{2}\int_0^1 \frac{(d\chi(s))^2}{ds}\right]\prod_{s\neq 0}\frac{\delta\chi(s)}{\sqrt{2\pi ds}} \; ;$$

or, equivalently, as the integral

$$\int^{\infty} f[\phi]\exp\left[-\int_0^1 \phi^2(s)ds\right]\prod_s \sqrt{\frac{ds}{2\pi}}\,\delta\phi(s)$$

of the function

$$f[\phi(\cdot)] = h\left[\int_0^{\cdot} \phi(s)ds\right] \quad .$$

The manner in which the Wiener integral is defined differs from the manner in which we had previously defined the integral of a functional $f[\cdot]$. The Wiener integral is not defined with the aid of projections but as the Lebesgue integral with respect to the appropriate totally additive measure. The space of functions ϕ -- or of functions χ must, therefore, be taken so wide that the measure becomes totally additive. This would be the case if all functions χ were admitted, no restriction being imposed on their values -- except $\chi(0) = 0$. But it is not necessary to go so far. Wiener had made the remarkable discovery that it is sufficient for this purpose to take the class of continuous functions χ. In fact, it was found to be sufficient to take the class of Hölder continuous functions with a Hölder exponent $\mu < \frac{1}{2}$. (Wiener employs $\mu = \frac{1}{4}$.)

Our intention is to give a simple derivation of this fact, without using the "transference" principle and without employing a Fourier transformation.

In the following we shall not be concerned with the integral but only with the measure; at first, only with the measure of cylinders.

To describe a cylinder we may introduce the projection of a function $\phi(\cdot)$ into a step function which is piecewise constant in the intervals $\mathcal{J}_\nu$ $(s_{\nu-1} < s < s_\nu)$ of a partition of the interval $0 < s < 1$. We assume $s_0 = 0$, $s_n = 1$. The value ϕ_ν of the projection in the interval $\mathcal{J}_\nu$ is given by

$$\phi_\nu = \int_{\mathcal{J}_\nu} \phi(s)ds \Big/ \Delta_\nu$$

with $$\Delta_\nu = (s_\nu - s_{\nu-1}) .$$

Evidently, we may describe this projection in terms of the function χ; namely, as the piecewise linear function which at the points $s_0 = 0, s_1, \ldots, s_n = 1$ has the same values $\chi_\nu = \chi(s_\nu)$ as the function χ.

A cylinder Cy may now be described as a set of functions ϕ -- or χ -- whose projected images lie in a given set of the $\{\phi_1, \ldots, \phi_n\}$ -- or $\{\chi_1, \ldots, \chi_n\}$ - space. The set of images is called the base BCy of the cylinder Cy.

The measure of such a cylinder Cy is then the measure of its base, which equals

$$\int_{BCy} \exp\left[-\tfrac{1}{2}\sum_\nu \phi_\nu^2 \Delta_\nu\right] \prod_\nu \sqrt{\frac{\Delta_\nu}{2\pi}}\, d\phi_\nu .$$

So far we have not named this class of functions from which these sets are taken. We now choose a positive number μ to be kept fixed throughout, $\mu < \frac{1}{2}$. Then we introduce the

space $\mathfrak{H}$ consisting of those functions χ to which there is associated a constant c such that the inequality

$$(*) \qquad |\Delta\chi| \leq c|\Delta s|^{\mu}$$

holds for any pair of points $s_1, s_2 \in [0,1]$ where

$$\Delta\chi = \left[\chi(s_2) - \chi(s_1)\right], \quad \Delta s = (s_2 - s_1) .$$

Such functions will be called "Hölder functions". The sets considered in the following will then consist of Hölder functions.

The statement to be proved is this: <u>the measures of cylinders</u> Cy <u>in the space</u> $\mathfrak{H}$ <u>are totally additive</u>. Any nested sequence of cylinders $Cy^1 \supset Cy^2 \supset \ldots$ in $\mathfrak{H}$ whose measures have a positive lower bound have at least one function $\chi(\cdot)$ in common.

We shall prove the <u>total additivity of the measure</u> first for a special class of cylinders, namely for <u>cylinders associated with partitions whose vertices are binary rationals</u> . We shall call such cylinders "binary Hölder cylinders" .

Specifically, we introduce the vertices

$$s_{n,\kappa} = 2^{-n}\kappa \quad (\kappa = 0,1,\ldots,2^n)$$

and the intervals

$$\mathcal{I}_{n,\kappa} = \left\{s : s_{n,\kappa-1} < s < s_{n,\kappa}\right\} (\kappa = 1,\ldots,2^n) .$$

Then we introduce the space $\mathfrak{G}_n$ of functions ϕ which are piecewise constant in the intervals $\mathcal{I}_{n,\kappa}$. By P_n we denote the projection which projects a function ϕ into $\mathfrak{G}_n$ by

$$P_n\phi(s) = 2^{-n} \int_{\mathcal{I}_{n,\kappa}} \phi(s')ds' \quad [s \in \mathcal{I}_{n,\kappa} (\kappa = 1,2,\ldots 2^n)] .$$

We employ the notation P_n also for the projection of a function χ into the piecewise linear function $P_n\chi$ whose derivative is the projection of the derivative of χ into $\mathcal{C}_n$.

Next we introduce the space $\mathfrak{K}_{c,N}$ of all those Hölder functions χ for which the relations

$$(*) \qquad |\chi(s_{\nu,\kappa}) - \chi(s_{\nu,\kappa-1})| = |\phi_{\nu,\kappa}|2^{-\nu} \leq ce^{-\nu\mu}$$

hold for all values $\nu \leq N$. We then shall estimate the measure $m(\mathfrak{K}_{c,N})$ of this space. In fact, we shall show that this measure can be made arbitrarily close to 1 by choosing c large enough independently of the value of N.

In order to estimate the measure of $\mathfrak{K}_{c,N}$ we shall construct a subspace of $\mathfrak{K}_{c,N}$ which is a product space and whose measure therefore can be computed. To this end we introduce the Rademacher orthogonal functions $u_{n,\lambda}$ for $n = 0,1,2,\ldots$ and $\lambda = 1,\ldots,2^{n-1}$,

$$u_0(s) = 1$$

$$u_{n,\lambda}(s) = \begin{cases} 1 & s\in\Delta_{n,2\lambda-1} \\ -1 & s\in\Delta_{n,2\lambda} \\ 0 & \text{otherwise} \end{cases}.$$

Evidently,

$$\int_0^1 u_{n,\lambda}^2(s)ds = 2^{1-n}$$

and

$$\int_0^1 u_{n,\lambda}(s)\,u_{n',\lambda'}(s)ds = 0 \text{ unless } n = n',\ \lambda = \lambda'.$$

We need only consider functions ϕ which belong to one of the spaces $\mathcal{G}_n$. Any such function possesses the expansion

$$\phi(s) = \sum_{n,\lambda} 2^{n-1} \beta_{n,\lambda} \; u_{n,\lambda}(s)$$

with the coefficients

$$\beta_{n,\lambda} = \int_{I_{n,2\lambda-1}} \phi(s)ds - \int_{I_{n,2\lambda}} \phi(s)ds \quad , \quad n \neq 0 \, ,$$

$$\beta_0 = \int_0^1 \phi(s)ds \, .$$

The integral functions χ then possess expansions

$$(**) \qquad \chi(s) = \sum_{n,\lambda} 2^{n-1} \beta_{n,\lambda} \; v_{n,\lambda}(s)$$

in which $v_{n,\lambda}(s)$ is the piecewise linear function

$$v_{n,\lambda}(s) = \int_0^s u_{n,\lambda}(s')ds'$$

The coefficients $\beta_{n,\lambda}$ can be expressed in terms of $\chi(\cdot)$ by

$$\beta_{n,\lambda} = \left[2\chi(s_{n,2\lambda-1}) - \chi(s_{n,2\lambda}) - \chi(s_{n,2\lambda-2}) \right]$$

We now consider the space $\mathcal{R}_{\theta,N}$ of all functions χ in $\mathfrak{X}$ whose projections in $\mathcal{G}_N$ possess expansion coefficients $\beta_{n,\lambda}$ which satisfy the conditions

$$(***) \qquad |\beta_{n,\lambda}| \leq \theta 2^{-\mu n} \qquad \text{for each } n \leq N \, .$$

This space is evidently a cylinder ; its measure can easily be evaluated. Before doing this we shall show that if

$$\theta = (2-2^{\mu})c$$

then every function in $\mathfrak{R}_{\theta,N}$ lies in $\mathfrak{F}_{c,N}$.

To this end we consider any function in $\mathfrak{S}_N$ and estimate the differences $|\chi(s_{\nu,\kappa}) - \chi(s_{\nu\kappa-1})|$ which enter the condition (*) through which the space $\mathfrak{F}_{c,N}$ was defined. We employ the representation (**) of χ. Note that the function $v_{n,\lambda}$ vanishes except in the interval $\Delta_{n-1,\lambda}$. For any binary point $s = s_{\nu,\kappa}$, therefore, the terms with $n > \nu$ drop out from the series (**). For each value of $n \leq \nu$ there is evidently <u>exactly one</u> value of λ for which $|v_{n,\lambda}(s_{\nu,\kappa}) - v_{n,\lambda}(s_{\nu,\kappa-1})| \neq 0$. For <u>this</u> value of λ we have

$$|v_{n,\lambda}(s_{\nu,\kappa}) - v_{n,\lambda}(s_{\nu,\kappa-1})| = [s_{\nu,\kappa} - s_{\nu,\kappa-1}] = 2^{-\nu}.$$

Therefore we may deduce from (**) the estimate

$$|\chi(s_{\nu,\kappa}) - \chi(s_{\nu,\kappa-1})| \leq \sum_{n=0}^{\nu} 2^{n-1}\theta 2^{-\mu n} 2^{-\nu}$$

$$\leq 2^{-1-\nu}\theta[2^{(1-\mu)}-1]^{-1}\, 2^{(1-\mu)(\nu+1)}$$

$$= c2^{-\mu\nu} \quad .$$

Thus inequality (*) is established and it follows that $\mathfrak{R}_{\theta,N}$ is contained in $\mathfrak{F}_{c,N}$.

As a consequence we have the relation

$$m(\mathfrak{F}_{c,N}) \geq m(\mathfrak{R}_{\theta,N})$$

for the measures of these spaces.

In order to evaluate the measure of the space $\mathcal{R}_{e,N}$ we note that

$$\int_0^1 \phi^2(s)ds = \sum_{n,\lambda} 2^{n-1}\, \beta_{n,\lambda}^2$$

for each function ϕ in $\mathcal{C}_N$, as is easily verified. Therefore, the measure of $\mathcal{R}_{\theta,N}$ is given by the integral

$$m(\mathcal{R}_{\theta,N}) = \int\cdots\int_{(***)} \exp\left[-\frac{1}{2}\sum_{n,\lambda} 2^{n-1}\,\beta_{n,\lambda}^2\right]\sqrt{\frac{2^{n-1}}{2\pi}}\prod_{n,\lambda} d\beta_{n,\lambda}$$

where the integration is to be extended over the set of all $\beta_{n,\lambda}$ satisfying inequalities (***). Since these inequalities restrict each coefficient separately we have

$$m(\mathcal{R}_{\theta,N}) = \prod_{\substack{n,\lambda \\ N\le n}} \int_{|\beta_{n,\lambda}|\le \theta 2^{-\mu n}} \exp\left[-\frac{1}{2}\,2^{n-1}\,\beta_{n,\lambda}^2\right]\sqrt{\frac{2^{n-1}}{2\pi}}\,d\beta_{n,\lambda} .$$

We introduce the function

$$\Phi(z) = \frac{1}{\sqrt{2\pi}}\int_{-z}^{z} \exp[-u^2/2]du$$

and set

$$\rho = \frac{1}{2} - \mu .$$

Since there are 2^{n-1} values of λ associated with the subscript n we have

$$m(\mathcal{R}_{\theta,N}) = \prod_{n\le N}\left[\Phi\left(2^{\rho n}\theta/\sqrt{2}\right)\right]^{2^{n-1}}$$

Now we use the assumption, $\mu < \frac{1}{2}$, made when the Hölder exponent μ was introduced. This condition implies $\rho > 0$ and

and hence we see that the argument, $2^{\rho n}\theta/\sqrt{2}$, of the function $\Phi(z)$ tends to infinity with n.

We note that, for any $z > 0$,

$$-\exp[-z^2/2]/z = \int_z^\infty \frac{d}{du}(\exp[-u^2/2]/u)du$$

$$= -\int_z^\infty (\exp[-u^2/2]/u^2)du - \int_z^\infty \exp[-u^2/2]du$$

$$= -\int_z^\infty \exp[-u^2/2]du \left[1+\mathcal{O}\left(\frac{1}{z^2}\right)\right]$$

from which it follows that

$$\Phi(z) = 1-\frac{2}{\sqrt{2\pi}}\int_z^\infty \exp[-u^2/2]du$$

$$= 1-\left(\sqrt{\frac{2}{\pi}}\exp[-z^2/2]/z\right)\Big/\left[1+\mathcal{O}\left(\frac{1}{z^2}\right)\right]$$

and so there are numbers $\theta_0 > 0$ and $\gamma > 0$ such that, for $z > \theta_0/\sqrt{2}$,

$$\Phi(z) \geq \exp\left[-\gamma\sqrt{2}\exp[-z^2/2]\right].$$

Hence, for $\theta \geq \theta_0$,

$$m(R_{\theta,N}) \geq \exp\left[\sum_{n=1}^{N} 2^{n-1}\left(-\gamma\sqrt{2}\exp[-2^{2\rho n}\theta^2/4]/[2^{\rho n}\theta/\sqrt{2}]\right)\right]$$

$$= \exp\left[-\gamma\theta^{-1}\sum_{n=1}^{N}2^{n(1-\rho)}\exp[-4^{\rho n-1}\theta^2]\right]$$

$$\geq \exp[-\Gamma/\theta]$$

where

$$\Gamma = \gamma \sum_{n=1}^{\infty} 2^{n(1-\rho)} \exp[-4^{\rho n-1}\theta_0^2];$$

evidently, as $\mu < \frac{1}{2}$, $\rho > 0$ so $\Gamma < \infty$.

At the same time we have, with $z = (2-2^{\mu})\Gamma$, $c_0 = \theta_0/(2-2^{\mu})$,

$$m(\mathfrak{H}_{c,N}) \geq e^{-z/c} \quad \text{for } c \geq c_0 .$$

Having established this estimate we proceed to prove the total additivity of the measure of finary Hölder cylinders. In this connection it is convenient to introduce the class of functions $\hat{x}$ which are defined only for binary values of s. Such functions will be referred to as "binary functions" and the space of these functions, without restriction of their values, will be denoted by $\hat{\mathfrak{B}}$. The subspace of functions $\hat{x}$ whose values satisfy the condition

$$(*) \qquad |\hat{x}(s_{\nu,\kappa}) - \hat{x}(s_{\nu,\kappa-1})| \leq c2^{-\nu}$$

for $\nu \leq N$ will be denoted by $\hat{\mathfrak{H}}_{c,N}$; the subspace of those for which (*) is satisfied for <u>all</u> values of ν will be denoted by $\hat{\mathfrak{H}}_c$; and the space of those which belong to $\hat{\mathfrak{H}}_c$ for some c will be denoted by $\hat{\mathfrak{H}}$. Such binary functions will be called "binary Hölder functions".

Every Hölder function evidently "generates" a binary Hölder function and, vice versa, for every binary Hölder functions there is a unique Hölder function -- with the same value of c -- which generates it. For this reason we need not distinguish between Hölder functions as binary Hölder functions. We express this fact, symbolically, by

$$\hat{\mathfrak{H}}_c = \mathfrak{H}_c, \qquad \hat{\mathfrak{H}} = \mathfrak{H}$$

On the other hand, the space $\hat{\mathfrak{H}}_{c,N}$ contains binary functions which are not Hölder functions. We express this fact by

$$\hat{\mathfrak{H}}_{c,N} \supset \mathfrak{H}_{c,N} \quad .$$

Furthermore it is clear that $\hat{\mathfrak{H}}_{\mu,c}$ is the common intersection of all spaces $\hat{\mathfrak{H}}_{\mu,c,N}$

$$(+) \qquad \hat{\mathfrak{H}}_{c} = \bigcap_{N} \hat{\mathfrak{H}}_{c,N} \quad ,$$

and

$$(++) \qquad \mathfrak{H} = \bigcup_{c} \hat{\mathfrak{H}}_{c} \quad .$$

We now consider cylinders $\hat{\mathfrak{C}}$ composed of binary functions, called "binary cylinders" and the Borel sets, called "binary Borel sets", generated by these binary cylinders. Evidently, $\mathfrak{H}_{c}$ is such a binary Borel set.

To each binary cylinder $\hat{\mathfrak{C}}$ we assign the binary Hölder cylinder $\mathfrak{C}$ in $\mathfrak{H}$, i.e., the cylinder composed of Hölder functions which has the same base. Then

$$\mathfrak{C} = \hat{\mathfrak{C}} \cap \hat{\mathfrak{H}}$$

Hence $\mathfrak{C}$ is a binary Borel set.

To every binary cylinder $\hat{\mathfrak{C}}$ we assign as its measure $m(\hat{\mathfrak{C}})$ the measure of its base; thus

$$\hat{m}(\hat{\mathfrak{C}}) = m(\mathfrak{C}) \quad .$$

In particular we have

$$\hat{m}(\hat{\mathfrak{H}}_{c,N}) = m(\mathfrak{H}_{c,N}) \quad .$$

The inequality established for the latter measure on page X-7 thus yields the relation

$$(+++) \qquad \hat{m}(\hat{\mathfrak{K}}_{c,N}) \geq e^{-z/c} \qquad \text{for } c \geq c_0 \ .$$

Now we maintain that this measure is totally additive for binary cylinders. This fact is implied by Kolmogoroff's main theorem*. One of the consequences of this fact is that relation (+) yields the inequality

$$\hat{m}(\hat{\mathfrak{K}}_c) \geq \text{l.u.b.}_N \ \hat{m}(\mathfrak{K}_{c,N}) \ .$$

Since the estimate (+++) is independent of N we have

$$\hat{m}(\mathfrak{K}_c) \geq e^{-z/c} \qquad \text{for } c \geq c_0 \ .$$

Hence (++) yields

$$\hat{m}(\hat{\mathfrak{K}}) = 1 \ .$$

In other words, within the space of unrestricted binary functions the space of Hölder functions has the measure 1 .

From $\mathfrak{C} = \hat{\mathfrak{C}} \cap \hat{\mathfrak{K}}_0$ we now conclude that the measure m is totally additive on the Hölder cylinders $\mathfrak{C}$. At the same time we conclude

$$\hat{m}(\mathfrak{C}) = \hat{m}(\hat{\mathfrak{C}}) = m(\mathfrak{C}) \ .$$

---

*See the "Extension Theorem", p.17, Foundations of Probability Theory by A.N. Kolmogoroff, Chelsea 1956.

Thus it follows that the measure m, originally assigned to Hölder cylinders, is totally additive on the binary Hölder cylinders. For the validity of this statement it is, of course, irrevelant whether or not we regard the Hölder functions composing $\mathfrak{C}$ represented by their values at binary points.

We now must consider general -- not necessarily binary -- cylinders composed of Hölder functions. We maintain that any Hölder cylinder in $\mathfrak{H}$ is a Borel set generated by binary Hölder cylinders provided that its base is a Borel set generated by rectangular cells. It is obviously no restriction to assume that the base of the cylinder $\mathfrak{C}$ is an open rectangular cell, given by inequalities of the form

$$a_a < \chi(s_a) < b_a \quad , \quad a = 1, \ldots m \quad .$$

In fact, it is no restriction to assume, in addition, that the functions χ have a bounded Hölder constant, c,

$$(\vartheta) \qquad |\chi(s') - \chi(s'')| \leq c|s' - s''|^{\mu} \quad ;$$

for cylinders without such a restriction are the limits of cylinders satisfying it for $c \longrightarrow \infty$.

Now let s_a^n be binary points, associated with the projection P_n, which tend to the points s_a as $n \longrightarrow \infty$, and let $\mathfrak{C}^n$ be the binary cylinder $\mathfrak{C}^n$ defined by the conditions

$$\left[a_a + c|s_a^n - s_a|^{\mu}\right] < \chi(s_a^n) < \left[b_a - c|s_a^n - s_a|^{\mu}\right].$$

Then we maintain that the cylinders $\mathfrak{C}^n$ tend to the cylinders $\mathfrak{C}$ in the set-theoretical sense. In fact, let the function χ be in $\mathfrak{C}$; since it satisfies the Hölder condition (ϑ) it will be in each $\mathfrak{C}^n$ from a certain n on, namely if n is chosen so large that

$$c|s_a^n - s_a|^{\mu} < \min\left\{(\chi(s_a) - a_a), (b_a - \chi(s_a))\right\} \quad .$$

If the function χ lies in infinitely many $\mathfrak{C}^n$ it is also in $\mathfrak{C}^\infty$, again by virtue of the Hölder inequality. Thus it is seen that the Hölder cylinders $\mathfrak{C}^\infty$ are Borel sets generated by binary cylinders.

The measure which was assigned to a cylinder, namely the Gaussian measure of its base, is evidently the limit of the measure of the approximating binary cylinders. It is a corollary of the theory of Borel measure that the measure of Borel sets derived in this fashion from the measures of the generating sets is totally additive if the measure of the generating set has this property.

Thus it is proved that the Wiener measure assigned to cylinders composed of Hölder functions with a fixed exponent $\mu < \frac{1}{2}$ is totally additive. The total additivity of m over the continuous function is achieved by assigning the measure zero to the set of continuous functions not in $\mathfrak{C}_\mu$.

Chapter XI

Stochastic Processes

A "stochastic process", which we shall denote by $[X_t, T, \Omega, \beta, p]$, consists of a space Ω, a Borel field β of subsets of Ω together with a countably additive, non-negative measure p on β such that $p(\Omega) = 1$, a parameter set T, and, for each $t \in T$, a real-valued measurable function (called a "random variable") $X_t(.)$ defined on Ω.

Thus, to each $t \in T$ there corresponds a function X_t. We may consider, for any fixed $\omega_o \in \Omega$, the values, for each $t \in T$, of $X_t(\omega_o)$ as a function (called a "sample function") defined on T. Thus, for each $\omega \in \Omega$ there is a function defined by

$$f_\omega(t) = X_t(\omega) \qquad \text{for } t \in T.$$

The space of sample functions can be made a measure space in the obvious way.

In many applications one is not given β and p but only $\{X_t\}, T$, and Ω, together with some arbitrary "probability distributions" (see below) for which it is desirable to construct β and p in some consistent way. In such a situation one usually takes for β the smallest Borel field such that all the random variables $\{X_t\}$ are measurable.

As an example let us consider a stochastic process of what is called "function space type". Let Ω be a space of functions with domain $[0,\infty)$, let T be $[0,\infty)$, and take the random variables to be defined by

$$X_t(\omega) = \omega(t)$$

for $t \in [0,\infty) = T$ and $\omega = \omega(\cdot) \in \Omega$. Note that in this case the space of sample functions is identical with the underlying space Ω. We shall define β as follows. Let β_o be the collection of all subsets of Ω of the form

$$\left\{\omega:\ \omega \in \Omega,\ [X_{t_1}(\omega),\ldots X_{t_n}(\omega)] \in A\right\}$$

$$=\left\{\omega :\ \omega = \omega(\cdot) \in \Omega,\ [\omega(t_1),\ldots,\omega(t_n)] \in A\right\}$$

where $t_1 \ldots t_n$ are n fixed numbers in $[0,\infty)$ and A is an n-dimensional Borel set. Then let β be the smallest Borel field containing β_o.

This is still not a stochastic process until the measure p is specified. We shall make use of two theorems of Kolmogoroff which will enable us to find measures in a fairly easy way. First, however, let us define a probability distribution. We say that a function $F = F(\lambda_1,\ldots,\lambda_n)$ defined on real Euclidean n-space is an n-dimensional distribution function if

1) F is monotonic non-decreasing and left-continuous in each variable independently,

2) $\lim_{\lambda_j \to -\infty} F(\lambda_1^*,\ldots\lambda_j\ldots\lambda_n^*) = 0$ for $j = 1\ldots n$ and arbitrary fixed finite $\left\{\lambda_i^*\right\}$, $i \neq j$, and

3) $\lim_{\lambda_1 \to \infty,\ldots\lambda_n \to \infty} F(\lambda_1\ldots\lambda_n) = 1$.

The first theorem is the Kolmogoroff Extension Theorem.

Theorem I: If a collection β_o of sets generates a Borel field β and if p_o is a countably additive non-negative set-function on β_o then there is a uniquely determined countably additive non-negative measure p on β coinciding with ε_o on β_o. We call p the extension of p_o to β.

The second theorem shows how to construct, using the notion of distribution function, such a set-function p_o without having to verify directly its countable additivity. (NB. We state a special case of this theorem applicable to our example.)

Theorem II: To every collection $\left\{t_1\ldots t_n\right\}$ of n parameter values (n = 1,2...) let us assign an n-dimensional distribution function $F_{t_1\ldots t_n}(\cdot)$ satisfying the following compatibility conditions -

1) If $\{i_1 \ldots i_n\}$ is any permutation of the integers $\{1 \ldots n\}$ and if $\{\lambda_1 \ldots \lambda_n\}$ is any point in the argument space (λ_i real) then

$$F_{t_1 \ldots t_n}(\lambda_1 \ldots \lambda_n) = F_{t_{i_1} \ldots t_{i_n}}(\lambda_{i_1} \ldots \lambda_{i_n}) \quad \text{and}$$

2) if $1 \leq k < n$ then, writing

$$\lim_{\lambda_{k+1} \to \infty \ldots \lambda_n \to \infty} F_{t_1 \ldots t_n}(\lambda_1 \ldots \lambda_n) = F_{t_1 \ldots t_n}(\lambda_1 \ldots \lambda_k, \infty, \ldots \infty),$$

we must have

$$F_{t_1 \ldots t_n}(\lambda_1 \ldots \lambda_k, \infty \ldots \infty) = F_{t_1 \ldots t_k}(\lambda_1 \ldots \lambda_k).$$

Then every set of β_o is of the form

$$B = \left\{\omega : \omega = \omega(\cdot) \in \Omega, [\omega(t_1) \ldots \omega(t_n)] \in A\right\}$$

for some $A \leq E_n$ and we define $p_o(B)$ as

$$p_o(B) = \int_A^n dF_{t_1 \ldots t_n}(\lambda_1 \ldots \lambda_n) .$$

Theorem II, then, states that the set-function p_o on β_o so constructed is countably additive. It is obviously non-negative and so satisfies the condition of Theorem I.

We thus see that for a process of function space type one can define a stochastic process given the space Ω, the parameter set T, the random variables $\{X_t\}$, and a compatible specification of the distributions of each finite collection of random variables. Although we have shown this only for processes of function space type we note that any abstract process can be represented by one of this type by means of the sample functions $\{x_\omega\}$.

Before proceeding to a familiar example — the Brownian Motion Process — let us insert a couple of definitions.

We say that a stochastic process $\{X_t, T, \Omega, \beta, P\}$ has "independent increments" if, for $t_o < t_1 < \ldots < t_n$ ($t_i \in T$, real) and 1-dimensional Borel sets A_i ($i = 1, \ldots n$),

$$p\left(\left\{\omega : \omega \in \Omega,\ [X_{t_i}(\omega) - X_{t_{i-1}}(\omega)] \in A_i\ ,\quad i = 1,\ldots n\right\}\right)$$

$$= \prod_{i=1}^{n} p\left(\omega : \omega \in \Omega,\ [X_{t_i}(\omega) - X_{t_{i-1}}(\omega)] \in A_i\right\}\right) .$$

We say that the stochastic process has "orthogonal increments" if, for $t_1 < t_2 \leq s_1 < s_2$ $(t_i, s_i \in T$, real),

$$\int_{\Omega} [X_{t_2}(\omega) - X_{t_1}(\omega)][X_{s_2}(\omega) - X_{s_1}(\omega)]p(d\omega) = 0 .$$

Note that a stochastic process with independent or orthogonal increments is usually obtained by specifying the distributions of the increment rather than those of the random variables themselves. To make the original functions random variables one takes $X_o(\omega) \equiv 0$ so that each X_t is itself an increment.

We now consider the stochastic process corresponding to the physical phenomenon of Brownian Motion. If microscopic particles are observed in a fluid they are seen to move in an irregular manner due to the imbalances of the impacts of the molecules of the fluid. To a first approximation the X-coordinate of the particle at time $(t_o + t)$ defines a real-valued function for $t \in [0,\infty)$ which may be identified with a sample function $\chi(t)$ of a stochastic process of function space type. The underlying space Ω, as was noted above, may be identified with the space of sample functions and consists of the functions $\{\omega(\cdot)\}$ defined on $[0,\infty)$ and such that $\omega(0) = 0$. As before $T = [0,\infty)$ and $X_t(\omega) = \omega(t)$. We define β_o as before and, letting $t_o = 0$, set

$$F_{t_1\ldots t_n}(\lambda_1\ldots\lambda_n)$$

$$= (2\pi)^{-\frac{n}{2}} \left[\prod_{i=1}^{n} (t_i - t_{i-1})\right]^{-\frac{1}{2}} \int_{-\infty}^{\lambda_1} \cdots \int_{-\infty}^{\lambda_n} e^{-\left[\frac{1}{2}\sum_{i=1}^{n} \frac{u_i^2}{t_i - t_{i-1}}\right]} du_1 \ldots du_n .$$

These are clearly compatible distribution functions. From them we define p_o by Theorem II and then the measure p on β by Theorem I. This measure p is known as the Wiener measure on the space of real-valued functions on $[0,\infty)$. This stochastic process has independent orthogonal increments and

$$\int_{\Omega}[X_t(\omega) - X_s(\omega)]^2 p(d\omega) = \int_{\Omega}[\omega(t) - \omega(s)]^2 p(d\omega) = |t-s|$$

which implies that for almost all $\omega \in \Omega$

$$\omega(t) \to \omega(s) \qquad \text{as } t \to s$$

so the continuous functions in Ω have measure 1 (almost all functions in Ω are continuous).

We consider, now, a general stochastic process (say of function space type) for which the parameter set T is non-denumerable (say $[0,\infty)$). One would like to be able to determine, for example, the probability that a function $\omega \in \Omega$ be bounded on a given interval $\mathcal{I}$ by a given constant C; i.e.

$$p\left(\left\{\omega : \underset{t\in\mathcal{I}}{\text{l.u.b.}} |\omega(t)| \leq C\right\}\right).$$

Unfortunately, this question has no meaning as the set $\left\{\omega : \underset{t\in\mathcal{I}}{\text{l.u.b.}} |\omega(t)| \leq C\right\}$ is not necessarily measurable. In fact, in general sets defined by a relation as above on a set which is non-denumerable will not be in β. [We shall prove that, if $\mathcal{I}$ is non-denumerable set, $\left\{\omega : \underset{t\in\mathcal{I}}{\text{l.u.b.}} |\omega(t)| \leq C\right\}$ has inner measure 0.] We might extend β to include such sets but their measure would not be determined by the finite dimensional distributions. One way out of this is the introduction of the notion of separability.

Before doing so let us prove the

Theorem III: If $\mathcal{I}$ is non-denumerable then the set $\wedge$, where

$$\wedge = \left\{\omega : \underset{t\in\mathcal{I}}{\text{l.u.b.}} |\omega(t)| < C\right\},$$

has inner measure zero.

Proof: Let $\bigwedge'$ be the complement of $\bigwedge$ and let $\Gamma = \bigcup_{i=1}^{\infty} \Gamma_i$ be a covering of $\bigwedge'$ with each $\Gamma_i \in \beta_o$. As each $\Gamma_i \in \beta_o$ it is defined by relationships involving only a finite set of parameters. Altogether, then, there is denumerable set $\{t_j\} \subset T$ such that each Γ_i is determined by relations involving only parameter values in $\{t_j\}$. Let $\omega_o = \omega_o(\cdot)$ be any function in Ω. Then we can construct a function $\omega_1 \in \Omega$ such that

$$\omega_1(t_j) = \omega_o(t_j) \qquad (j = 1,2,\ldots)$$

and, for some $s \in \mathcal{S}$, $s \notin \{t_j\}$,

$$|\omega_1(s)| > C.$$

Then $\omega_1 \in \bigwedge' \subseteq \Gamma$ and so, as the values of ω_o on $\{t_j\}$ are identical with those of ω_1 and these alone are sufficient to determine membership in Γ, $\omega_o \in \Gamma$. This proves $\Gamma = \Omega$, $p(\Gamma) = 1$ and, as this is true for any Γ of this type, the inner measure of $\bigwedge$ is zero.

Let us now, return to the notion of separability. We say that a stochastic process $[X_t, T, \Omega, \beta, p]$ with $T = [0, \infty)$ is separable if there exists a denumerable set $\{t_j\} \subset T$ and a set $\bigwedge \in \beta$ with $p(\bigwedge) = 0$ such that, for every closed interval A and every open interval $\mathcal{S}$,

$$[(S \cup S') - (S \cap S')] \subseteq \bigwedge$$

where

$$S = \{\omega: \text{for every } t \in \mathcal{S},\ X_t(\omega) \in A\}$$

and

$$S' = \{\omega: \text{for every } t \in [\{t_j\} \cap \mathcal{S}],\ X_t(\omega) \in A\}.$$

In essence this enables us, given a relation specified on an open set, to replace it by a relation on a denumerable set — except for at most a set of functions of measure zero.

We shall assume from here on that p is complete. Then every subset of $\bigwedge$ is itself measurable and has measure zero.

We note that, for a separable process and a complete measure, we have, given any open set $\mathcal{J}$, for almost all ω

$$(1)\qquad \begin{cases} \underset{t\in\mathcal{J}}{\text{g.l.b.}}\ X_t(\omega) = \underset{t_j\in\mathcal{J}}{\text{g.l.b.}}\ X_{t_j}(\omega) \\ \underset{t\in\mathcal{J}}{\text{l.u.b.}}\ X_t(\omega) = \underset{t_j\in\mathcal{J}}{\text{l.u.b.}}\ X_{t_j}(\omega) \ . \end{cases}$$

Thus, $f(\omega) = \underset{t\in\mathcal{J}}{\text{l.u.b.}}\ X_t(\omega)$ and so on are random variables. The relations (1) could have been used to define separability.

Theorem IV: If there is a denumerable set $\{t_j\}\subset T$ and if for each open interval $\mathcal{J}$ there is a set $\bigwedge_{\mathcal{J}}\in\beta$ such that $p(\bigwedge_{\mathcal{J}}) = 0$ and (1) holds for all $\omega\notin\bigwedge_{\mathcal{J}}$, then the process is separable.

Proof: Enumerate all the open intervals with rational end-points calling them $\{\mathcal{J}_r\}$. Let $\bigwedge = \bigcup_{r=1}^{\infty}\bigwedge_{\mathcal{J}_r}$. Then

$$p(\bigwedge) \le \sum_{r=1}^{\infty} p(\bigwedge_{\mathcal{J}_r}) = 0 \ .$$

We note that, even for $\mathcal{J}\notin\{\mathcal{J}_r\}$, $\bigwedge_{\mathcal{J}}\subseteq\bigwedge$ as, for $\omega\notin\bigwedge$,

$$\underset{t\in\mathcal{J}}{\text{g.l.b.}}\ X_t(\omega) = \underset{\mathcal{J}_r\subset\mathcal{J}}{\text{g.l.b.}}\left\{\underset{t\in\mathcal{J}_r}{\text{g.l.b.}}\ X_t(\omega)\right\} = \underset{\mathcal{J}_r\subset\mathcal{J}}{\text{g.l.b.}}\left\{\underset{t_j\in\mathcal{J}_r}{\text{g.l.b.}}\ X_{t_j}(\omega)\right\}$$

$$= \underset{t_j\in\mathcal{J}}{\text{g.l.b.}}\ X_{t_j}(\omega) \ .$$

For any given A and $\mathcal{J}$, then, disregarding points of $\bigwedge$,

$$S = \left\{\omega:\left[\underset{t\in\mathcal{J}}{\text{g.l.b.}}\ X_t(\omega)\right]\in A, \left[\underset{t\in\mathcal{J}}{\text{l.u.b.}}\ X_t(\omega)\right]\in A\right\}$$

$$= \left\{\omega:\left[\underset{t_j\in\mathcal{J}}{\text{g.l.b.}}\ X_{t_j}(\omega)\right]\in A, \left[\underset{t_j\in\mathcal{J}}{\text{l.u.b.}}\ X_{t_j}(\omega)\right]\in A\right\} = S'$$

so the process is separable.

Note, also, that the inclusion of (a denumerable number of) additional points in $\{t_j\}$ leaves (1) unaltered.

We define the limit in the mean (abbreviated L.i.m.) of a sequence of random variables as follows.

L.i.m. $f_n(\omega) = f(\omega)$ means that $\int_\Omega |f_n(\omega)|^2 p(d\omega) < \infty$ for $n \to \infty$ every n and

$$\int_\Omega |f_n(\omega)-f(\omega)|^2 p(d\omega) \longrightarrow 0 \quad \text{as } n \longrightarrow \infty$$

This integral is defined as the usual Lebesgue - Stieltjes integral.

We now prove the

Theorem V: Let $[X_t, [0,\infty), \Omega, \beta, p]$ be a separable stochastic process such that, for any $t_o \in [0,\infty)$,

$$\underset{t \to t_o}{\text{L.i.m.}}\ X_t(\omega) = X_{t_o}(\omega) .$$

Further, let

$$a \le S_0^{(n)} \le \cdots \le S_{a_n}^{(n)} \le b$$

be a sequence of subdivisions of [a,b] becoming more and more dense as $n \to \infty$. Then

$$\underset{n \to \infty}{\text{L.i.m.}}\ \max_{0 \le j \le a_n} X_{S_j^{(n)}}(\omega) = \text{l.u.b.}_{t \in [a,b]}\ X_t(\omega)$$

for almost all $\omega \in \Omega$.

Proof: Write $X_t(\omega) = x(t,\omega)$. For any given t' choose, for each n, an integer j(n) so that

$$S_{j(n)}^{(n)} \longrightarrow t' \qquad \text{as } n \longrightarrow \infty$$

Then

$$\underset{n \to \infty}{\text{L.i.m.}}\ x(S_{j(n)}^{(n)}, \omega) = x(t', \omega) .$$

Then it is well known (see Doob, Stochastic Processes, pp. 8 and 9) that there is a subsequence $x(S^{(n_\nu)}_{j(n_\nu)}, \omega)$ which converges pointwise to $x(t',\omega)$ for almost all $\omega \in \Omega$. It follows that

$$x(t',\omega) \leq \lim_{n_\nu \to \infty} \inf \max_{0 \leq k \leq a_{n_\nu}} x(S^{(n_\nu)}_k, \omega) .$$

Certainly, however,

$$\lim_{n \to \infty} \inf \max_{0 \leq k \leq a_{n_\nu}} x(S^{(n_\nu)}_k, \omega) \leq \text{l.u.b.}_{t \in [a,b]} x(t,\omega)$$

which proves the theorem.

Happily, in the base space of separable process all interesting sets are measurable. This will not, however, be useful unless it can be shown that any stochastic process can be replaced in some reasonable way by a separable process. This can always be done.

We shall prove the

Theorem VI: Let $[X_t, [0,\infty), \Omega, \beta, p]$ be a stochastic process of function space type. Then there exists a set $\Omega_1 \subseteq \Omega$, not necessarily measurable but of outer measure 1 and there exists a denumerable set $\{t^*_j\}$ such that, for this $\{t^*_j\}$ and for all $\omega \in \Omega_1$, (1) holds.

Before proving this theorem we shall state and prove three lemmas.

Lemma (1): Let $\mathcal{J}$ be any open interval and let $\{S_j\}$ be any sequence in $\mathcal{J}$. Then there is a sequence $\{\bar{t}_j\}$ depending only on the interval $\mathcal{J}$ such that for almost all $\omega \in \Omega$

$$\begin{cases} \text{l.u.b.}_j [x(S_j,\omega)] \leq \text{l.u.b.}_j [x(\bar{t}_j,\omega)] \\ \text{g.l.b.}_j [x(S_j,\omega)] \geq \text{g.l.b.}_j [x(\bar{t}_j,\omega)] . \end{cases}$$

Lemma (2): There is a fixed sequence $\{t_j^*\}$ such that, for any open interval (or for that matter, any open set) $\mathcal{J}$ and for any sequence $\{S_j\}$ in $\mathcal{J}$,

$$(a^*) \qquad \begin{cases} \text{l.u.b.}_j\ [x(S_j,\omega)] \le \text{l.u.b.}_{t_j^* \in \mathcal{J}}\ [x(t_j^*,\omega)] \\ \text{g.l.b.}_j\ [x(S_j,\omega)] \ge \text{g.l.b.}_{t_j^* \in \mathcal{J}}\ [x(t_j^*,\omega)] \ . \end{cases}$$

Lemma (3): For any given t define

$$\begin{cases} x_*(t,\omega) = \lim_{\varepsilon \downarrow 0}\ \text{g.l.b.}_{|t_j^*-t|<\varepsilon}\ [x(t_j^*,\omega)] \\ x^*(t,\omega) = \lim_{\varepsilon \downarrow 0}\ \text{l.u.b.}_{|t_j^*-t|<\varepsilon}\ [x(t_j^*,\omega)] \ . \end{cases}$$

Then, for almost all $\omega \in \Omega$,

$$(1') \qquad x_*(t,\omega) \le x(t,\omega) \le x^*(t,\omega) \ .$$

Proof [of Lemma (1)]: Let

$$\begin{cases} M = \text{l.u.b.}_{\text{all sequences } \{\sigma_j\} \text{ in } \mathcal{J}} \int_\Omega \tan^{-1}(\text{l.u.b.}_j\ [x(\sigma_j,\omega)p(d\omega) \\ N = \text{g.l.b.}_{\text{all sequences } \{\sigma_j\} \text{ in } \mathcal{J}} \int_\Omega \tan^{-1}(\text{g.l.b.}_j\ [x(\sigma_j,\omega)])p(d\omega) \ . \end{cases}$$

Then there are sequences $\{S_j^{(k)}\}$ in $\mathcal{J}$ such that

$$\lim_{k\to\infty}\int_{\Omega}\tan^{-1}(\text{l.u.b.}_{j}\ [x(S_j^{(k)},\omega)])p(d\omega) = M\ .$$

Let $\{\hat{t}_j\}$ be the union of the sequences $\{S_j(k)\}$ arranged in some order. Then

$$\int_{\Omega}\tan^{-1}(\text{l.u.b.}_{j}\ [x(\hat{t}_j,\omega)])p(d\omega) = M\ .$$

Similarly, we define $\{\tilde{t}_j\}$ so that

$$\int_{\Omega}\tan^{-1}(\text{g.l.b.}_{j})[x(\tilde{t}_j,\omega)])\ p(d\omega) = N\ .$$

Finally, let $\{\bar{t}_j\}$ be the union of $\{\hat{t}_j\}$ and $\{\tilde{t}_j\}$ arranged in some order. We have, then, both

$$\begin{cases}\displaystyle\int_{\Omega}\tan^{-1}(\text{l.u.b.}_{j}\ [x(\bar{t}_j,\omega)])p(d\omega) = M \text{ and}\\ \displaystyle\int_{\Omega}\tan^{-1}(\text{g.l.b.}_{j}\ [x(\bar{t}_j,\omega)])p(d\omega) = N\ .\end{cases}$$

Let $\{s_j\}$ be any sequence in ϑ and suppose, for example, there were a set $\sum\in\beta$ with $p(\sum) > 0$ such that, for all $\omega\in\sum$,

$$\text{l.u.b.}_{j}\ [x(s_j,\omega) > \text{l.u.b.}_{j}\ [x(\bar{t}_j,\omega)]\ .$$

Let $\{s'_j\}$ be the union of $\{s_j\}$ and $\{\bar{t}_j\}$ ordered in some manner. Then, for all $\omega\in\Omega$,

$$\text{l.u.b.}_{j}\ [x(s'_j,\omega)] \geq \text{l.u.b.}_{j}\ [x(\bar{t}_j,\omega)]$$

and, for all $\omega\in\sum$,

$$\text{l.u.b.}_{j}\ [x(s'_j,\ )] > \text{l.u.b.}_{j}\ [x(\bar{t}_j,\omega)]$$

$$\text{l.u.b.}_j\ [x(s'_j,\omega)] > \text{l.u.b.}_j\ [x(\bar{t}_j,\omega)]$$

and, therefore,

$$\int_\Omega \tan^{-1}(\text{l.u.b.}_j[x(s'_j,\omega)])p(d\omega) > \int_\Omega \tan^{-1}(\text{l.u.b.}_j[x(\bar{t}_j,\omega)])p(d\omega) = M$$

contradicting the definition of M. Similarly, if there were a set $\sum'\varepsilon\beta$ with $p(\sum') > 0$ such that, for $\omega\varepsilon\sum'$,

$$\text{g.l.b.}_j\ [x(s_j,\omega)] < \text{g.l.b.}_j\ [x(\bar{t}_j,\omega)]\ ,$$

one would arrive at a contradiction of the definition of N. Thus, for every sequence $\{s_j\}$ in $\mathfrak{I}$ and the fixed sequence $\{\bar{t}_j\}$ in $\mathfrak{I}$, one has, for almost all $\omega\varepsilon\Omega$,

$$\begin{cases} \text{l.u.b.}_j\ [x(s_j,\omega)] \leq \text{l.u.b.}_j\ [x(\bar{t}_j,\omega)] \\ \\ \text{g.l.b.}_j\ [x(s_j,\omega)] \geq \text{g.l.b.}_j\ [x(\bar{t}_j,\omega)] \end{cases}$$

proving the lemma.

Proof [of Lemma (2)]: Let $\{\mathfrak{I}_r\}$ be the set of open intervals with rational endpoints in some order. Let $\{\bar{t}_{j,r}\}$ be the sequence $\{\bar{t}_j\}$ shown, by Lemma (1), to exist for $\mathfrak{I}_r$. Let $\{t^*_j\}$ be the union of all the $\{\bar{t}_{j,r}\}$ ordered in some manner. Clearly the inequalities (a^*) hold for any $\mathfrak{I}\varepsilon\{\mathfrak{I}_r\}$. Suppose in some open set $\mathfrak{I}$ there were a sequence $\{s_k\}$ such that, for example, on a set of positive measure

$$\text{l.u.b.}_k\ [x(s_k,\omega)] > \text{l.u.b.}_{t^*_j\varepsilon\mathfrak{I}}\ [x(t^*_j,\omega)]\ .$$

Then, for some k,

$$x(s_k,\omega) > \underset{t_j^*\varepsilon\mathcal{J}}{\text{l.u.b.}}\ [x(t_j^*,\omega)] \ .$$

As $\mathcal{J}$ is open we may find an open interval $\hat{\mathcal{J}}\varepsilon\{\mathcal{J}_r\}$ such that $s_k\varepsilon\hat{\mathcal{J}}\subset\mathcal{J}$ and we may take $\{\sigma_j\}$ to be any sequence in $\mathcal{J}$ such that $s_k\varepsilon\{\sigma_j\}$. Then

$$\underset{j}{\text{l.u.b.}}\ [x(\sigma_j,\omega)] \geq x(s_k,\omega) > \underset{t_j^*\varepsilon\mathcal{J}}{\text{l.u.b.}}\ [x(t_j^*,\omega)] \geq \underset{t_j^*\varepsilon\mathcal{J}}{\text{l.u.b.}}\ [x(t_j^*,\omega)]$$

contradicting (a^*) for $\hat{\mathcal{J}}\varepsilon\{\mathcal{J}_r\}$. Similarly one gets a contradiction by assuming the existence of an open set $\mathcal{J}'$ and a sequence $\{s_k\}$ in $\mathcal{J}'$ for which, on a set of positive measure

$$\underset{k}{\text{g.l.b.}}\ [x(s_k,\omega)] < \underset{t_j^*\varepsilon\mathcal{J}'}{\text{g.l.b.}}\ [x(t_j^*,\omega)].$$

Thus the lemma is proved.

Proof [of Lemma (3)]: Suppose there were a set $\sum\varepsilon\beta$ with $p(\sum) > 0$ such that, for $\omega\varepsilon\sum$,

$$x(t,\omega) < x_*(t,\omega) = \lim_{\varepsilon\downarrow 0}\ \underset{|t_j^*-t|<\varepsilon}{\text{g.l.b.}}\ [x(t_j^*,\omega)].$$

Then there would be an ε_0 such that, for $\omega\varepsilon\sum$ and $0 < \varepsilon < \varepsilon_0$,

$$\underset{|t_j^*-t|<\varepsilon}{\text{g.l.b.}}\ [x(t_j^*,\omega)] > x(t,\omega) \ .$$

That is, letting $\mathcal{J}$ be the open interval $(t-\varepsilon,t+\varepsilon)$ and letting $\{\tau_j\}$ be any sequence in $\mathcal{J}$ such that $t\varepsilon\{\tau_j\}$,

$$\text{g.l.b.}_{j}\ [x(\tau_j,\omega)] \le x(t,\omega) < \text{g.l.b.}_{t_j^*\varepsilon}\ [x(t_j^*,\omega)]$$

for $\omega\varepsilon\sum$ which contradicts Lemma (2). Similarly the assumption $x(t,\omega) > x^*(t,\omega)$ on a set of positive measure leads to a contradiction and the lemma is proved.

Proof [of the Theorem]: Let Λ_1 be defined by

$$\Lambda_1 = \{\omega: \text{for every } t\varepsilon[0,\infty),\ x_*(t,\omega) \le x^*(t,\omega)\}\ .$$

To show that Λ_1 has outer measure 1 let Γ be a countable β_o-covering of Λ_1 ; i.e. - let $\Gamma_i \epsilon \beta_o$ and $\Lambda_1 \subseteq \Gamma = \bigcup_{i=1}^{\infty} \Gamma_i$. Each of the Γ_i, being in β_o, is defined by relationships among the values of X_{t_k} for a finite number of t_k. Let $\{t_j\}$ be the union of all of these in some order.

Let

$$\Gamma_0 = \{\omega: \text{for } t\varepsilon\{\tau_j\},\ x_*(t,\omega) \le x(t,\omega) \le x^*(t,\omega)\}.$$

Then $p(\Gamma_0) = 1$ since its complement is the union of a denumerable number of null-sets. We have not, as yet, used the fact that the process is of function space type. Let $\omega(\cdot)\varepsilon\Lambda_1$ and let $\sigma(\cdot)$ be any element in Γ_0 . We write

$$\omega_*(t) = x(t,\omega) = \lim_{\varepsilon\downarrow 0}\ \text{g.l.b.}_{|t_j^*-t|<\varepsilon}\ [X_{t_j^*}(\omega)]$$

and so on. Then

$$\begin{cases} \omega_*(t) \le \omega(t) \le \overset{*}{\omega}(t) & \text{for all } t\ \varepsilon[0,\infty) \\ \sigma_*(t) \le \sigma(t) \le \overset{*}{\sigma}(t) & t\ \varepsilon\ \tau_j\ . \end{cases}$$

Let $\eta(t) = \begin{cases} \sigma(t) & t\in\{\tau_j\} \\ \omega(t) & t\notin\{\tau_j\} \end{cases}$.

Clearly $\eta(\cdot)\in\Omega_1$ and, therefore, $\eta(\cdot)\in\Gamma$. But, as $\sigma(\cdot)$ is equal, for $t\in\{\tau_j\}$, to a function in Γ and thus satisfies all the defining relationships of Γ, we have $\sigma(\cdot)\in\Gamma$. Hence $\Gamma_0\subseteq\Gamma$ and $p(\Gamma)=1$ so the outer measure of Ω_1 is 1. We must show that (1) holds for $\omega\in\Omega_1$. We know that (1') holds and must show that (1') implies (1). We note that

$$\text{g.l.b.}_{t\in\vartheta}\,[x_*(t,\omega)]$$

$$= \text{g.l.b.}_{t\in\vartheta}\ \lim_{\varepsilon\downarrow 0}\ \text{g.l.b.}_{|t_j^*-t|<\varepsilon}\,[x(t_j^*,\omega)]$$

$$= \text{g.l.b.}_{t_j^*\in\vartheta}\,[x(t_j^*,\omega)] \quad .$$

Then, from (1') ,

$$\text{g.l.b.}_{t\in\vartheta}\,[x(t,\omega)] \geq \text{g.l.b.}_{t\in\vartheta}\,[x_*(t,\omega)] = \text{g.l.b.}_{t_j^*\in\vartheta}\,[x(t_j^*,\omega)]$$

and as, clearly

$$\text{g.l.b.}_{t\in\vartheta}\,[x(t,\omega)] \leq \text{g.l.b.}_{t_j^*\in\vartheta}\,[x(t_j^*,\omega)]$$

we have

$$\text{g.l.b.}_{t\in\vartheta}\,[x(t,\omega)] = \text{g.l.b.}_{t_j^*\in\vartheta}\,[x(t_j^*,\omega)] \quad .$$

Similarly

$$\underset{t\varepsilon\}}{\text{l.u.b.}}\ [x(t,\omega)] = \underset{t_j^*\varepsilon\}}{\text{l.u.b.}}\ [x(t_j^*,\omega)]$$

which proves the theorem.

A new stochastic process $[X_t,[0,\infty),\Lambda,\beta_1,p_1]$ can now be formed where

$$\begin{cases} \beta_1 = \{S:\ S = (B_1\cap\Lambda_1)\cup(B_2\cap\Lambda_1^c);\ B_1,B_2\varepsilon\beta\} \\ p_1(S) = p(B_1) \end{cases}$$

This new process is separable and is an extension of the old as, on β, $p_1 = p$. The theory is still not quite satisfactory as the set Λ_1 is not uniquely determined.

We shall now apply the notion of separability to the study of Brownian motion processes.

We shall prove the

Theorem VII: Almost all of the sample functions of a separable Brownian motion process $[X_t,T,\Lambda,\beta,p]$ are continuous.

First we prove two lemmas.

Lemma (1): Let $\{y_k\}$ be a sequence of mutually independent random variables ($y_k\ \varepsilon\{X_t\}$) from any stochastic process $[X_t,T,\Lambda,\beta,p]$. That is, for every n and every n-tuple of Borel sets $\{B_k\}$,

$$p(\{\omega:y_k(\omega)\ \varepsilon B_k,\ k=1\ldots n\}) = \prod_{k=1}^{n} p(\{\omega:y_k(\omega)\ \varepsilon B_k\}).$$

Let $Y_n(\omega)=\sum_{k=1}^{n} y_k(\omega)$ and suppose that, for every m and n,

$$p(\{\omega:[Y_n(\omega)-Y_m(\omega)] < 0\}) = p(\{\omega:[Y_n(\omega)-Y_m(\omega)] > 0\}).$$

Then, for every n and every $\lambda > 0$,

$$p\left(\left\{\omega: \max_{0<j\leq n} [Y_j(\omega)] \geq \lambda\right\}\right) \leq 2\, p\left(\left\{\omega: Y_n(\omega) \geq \lambda\right\}\right).$$

Lemma (2): In a separable Brownian motion process, for any $a < b$ and any λ,

$$p\left(\left\{\omega: \left[\text{l.u.b.}_{a<t<b}\left\{x(t,\omega)\right\} - x(a,\omega)\right] \geq \lambda\right\}\right) \leq \frac{1}{\lambda}\sqrt{\frac{2(b-a)}{\pi}}\, e^{-\lambda^2/2(b-a)}.$$

Proof [of Lemma (1)]: Let $\nu = \nu(\omega)$ be such that, for $k < \nu$, $Y_k(\omega) < \lambda$ and $Y_\nu(\omega) \geq \lambda$ so that, for any ω, $Y_{\nu(\omega)}(\omega)$ is the first increment greater than or equal to λ. If for no finite k is $Y_k(\omega) \geq \lambda$ let $\nu(\omega) = \infty$. Using the hypotheses of independence and zero median,

$$p\left(\left\{\omega: \max_{j\leq n} [Y_j(\omega)] \geq \lambda \text{ and } Y_n < \lambda\right\}\right)$$

$$= \sum_{k=1}^{n-1} p\left(\left\{\omega: \nu(\omega) = k \text{ and } Y_n(\omega) < \lambda\right\}\right)$$

$$= \sum_{k=1}^{n-1} p\left(\left\{\omega: \nu(\omega) = k \text{ and } Y_n(\omega) < Y_k(\omega)\right\}\right)$$

$$= \sum_{k=1}^{n-1} p\left(\left\{\omega: \nu(\omega) = k\right\}\right) p\left(\left\{\omega: [Y_n(\omega) - Y_k(\omega)] < 0\right\}\right)$$

$$= \sum_{k=1}^{n-1} p\left(\left\{\omega: \nu(\omega) = k\right\}\right) p\left(\left\{\omega: [Y_n(\omega) - Y_k(\omega)] > 0\right\}\right)$$

$$= \sum_{k=1}^{n-1} p\left(\left\{\omega: \nu(\omega) = k \text{ and } Y_n(\omega) > Y_k(\omega) \geq \lambda\right\}\right)$$

$$\leq p\left(\left\{\omega: Y_n(\omega) \geq \lambda\right\}\right)$$

or

$$p\left(\left\{\omega: \max_{j\leq n} [Y_j(\omega)] \geq \lambda \text{ and } Y_n < \lambda\right\}\right) \leq p\left(\left\{\omega: Y_n(\omega) \geq \lambda\right\}\right).$$

Clearly

$$p(\{\omega: \max_{j\le n} [Y_j(\omega)] \ge \lambda \text{ and } Y_n \ge \lambda\}) \le p(\{\omega: Y_n(\omega) \ge \lambda\}) .$$

Hence, adding,

$$p(\{\omega: \max_{j\le n} [Y_j(\omega)] \ge \lambda\}) \le 2\, p(\{\omega: Y_n(\omega) \ge \lambda\}).$$

Proof [of Lemma (2)]: Let $\{t_0 \dots t_n\} \subset T = [0,\infty)$ be such that

$$a = t_0 < t_1 < \dots < t_n = b.$$

By Lemma (1),

$$p(\{\omega: \max_{j\le n} [x(t_j,\omega) - x(a,\omega) \ge \lambda\}) \le 2p(\{\omega: [x(b,\omega)-x(a,\omega)] \ge \lambda\}).$$

If we consider subdivisions

$$a = t_0^{(k)} < t_1^{(k)} < \dots < t_n^{(k)} = b$$

becoming more and more dense then, by Theorem IV,

$$p(\{\omega: \text{l.u.b.}_{a\le t\le b} [x(t,\omega)-x(a,\omega)] \ge \lambda\}) \le 2p(\{\omega: [x(b,\omega)-x(a,\omega)] \ge \lambda\})$$

$$= 2\,\frac{1}{\sqrt{2\pi(b-a)}} \int_\lambda^\infty e^{-\xi^2/2(b-a)} d\xi \le \frac{1}{\lambda}\sqrt{\frac{2(b-a)}{\pi}}\, e^{-\lambda^2/2(b-a)}.$$

Proof [of Theorem]: We shall actually prove that almost all $\omega(\cdot) \in \Omega$ are not only continuous but, in every bounded interval, uniformly continuous. Let us note that

$$p(\{\omega:\max_{j\le N^2}\ \text{l.u.b.}_{|Nt-j|<1}\ |x(t,\omega)-x(j/N,\omega)| \ge N^{-1/4}\})$$

$$\le \sum_{j=1}^{N^2} p(\{\omega:\ \text{l.u.b.}_{|Nt-j|<1}\ |x(t,\omega)-x(j/N,\omega)| \ge N^{-1/4}\})$$

$$\le \sum_{j=1}^{N^2} 2\sqrt{\frac{2}{\pi}}\ N^{-1/4} e^{-\sqrt{N/2}} = 2\sqrt{\frac{2}{\pi}}\ N^{7/4} e^{-\sqrt{N/2}}\ .$$

Noting that

$$\sum_{N=1}^{\infty} 2\sqrt{\frac{2}{\pi}}\ N^{7/4} e^{-\sqrt{N/2}} < \infty\ ,$$

we have, by the Borel-Cantelli lemma, the opposite inequality almost everywhere. That is, for every $N > 0$, we have, for almost all $\omega \in \Omega$, writing $x(\cdot,\omega) = \omega(\cdot)$,

$$(*) \qquad |\omega(t) - \omega(j/N)| \le N^{-1/4}$$

for all t, j such that

$$|t-j/N| < 1/N \quad \text{and} \quad t \in [0,N)\ .$$

Clearly, for almost every $\omega \in \Omega$, $(*)$ is true for every $N > 0$ since

$$p(\{\omega:\ \text{for some } N > 0,\ (*) \text{ is false}\})$$

$$\le \sum_{k=1}^{\infty} p(\{\omega:\ \text{for } N=k,\ (*) \text{ is false}\}) = 0\ .$$

Let $\bigwedge$ be the set of functions $\omega(\cdot)$ in Ω for which $(*)$ is true for all N. Note that $p(\bigwedge) = 1$. Then, for any $\omega \in \bigwedge$, given $\varepsilon > 0$ and $t \in [0,\infty)$ let $N > \max\left\{\frac{1}{2}\varepsilon^{-4},\ t\right\}$ and let $\delta = \frac{1}{N}$. For any t' such that $|t'-t| < \delta$ there is a $j \le N^2$ such that $|t-j/N| < 1/N$ and $|t'-j/N| < 1/N$. It follows that

$$|\omega(t) - \omega(t')| \le |\omega(t)-\omega(j/N)| + |\omega(t') - \omega(j/N)| \le 2N^{-1/4} < \varepsilon.$$

Hence each $\omega(\cdot) \in \bigwedge$ is uniformly continuous in every bounded subinterval of $T = [0,\infty)$.

We wish, now, to define a "stochastic integral". That is, if $[x(t,\cdot),[0,\infty),\Omega,\beta,p]$ is a stochastic process with orthogonal increments we shall try to give meaning to the expression

$$\int_0^\infty f(t)x(dt,\omega) .$$

We say that a stochastic process has orthogonal increments if, whenever $s < t \leq s' < t'$,

$$\int_\Omega [x(t,\omega)-x(s,\omega)][x(t',\omega)-x(s',\omega)]p(d\omega) = 0.$$

It follows (see Doob, Stochastic Processes, pp. 99-100) that there exists a unique non-negative measure dF [i.e. - a monotonic non-decreasing function $F(\cdot)$ unique to within an additive constant] such that

$$\int_\Omega [x(t,\omega)-x(s,\omega)]^2 p(d\omega) = F(t)-F(s) \text{ for } t > s \in [0,\infty).$$

Write

$$x^+(t,\cdot) = \underset{s\downarrow t}{\text{L.i.m.}}\ x(t,\cdot) \text{ and } x^-(t,\cdot) = \underset{s\downarrow t}{\text{L.i.m.}}\ x(t,\cdot) .$$

Then, except for an enumerable t-set,

$$x^+(t,\cdot) = x(t,\cdot) = x^-(t,\cdot) .$$

For the measure dF mentioned above let

$$\mathcal{L}_F^2 = \left\{f(\cdot): \int_0^\infty [f(t)]^2 dF(t) < \infty\right\}$$

and let

$$\mathcal{L}_\Omega^2 = \left\{\psi(\cdot): \int_\Omega [\psi(\omega)]^2 p(d\omega) < \infty\right\}$$

We will define, for any $f(\cdot) \in \mathcal{L}_F^2$, a functional $\hat{f}$ as

$$\hat{f}(\omega) = \int_0^\infty f(t)x(dt,\omega) \ .$$

We start by defining $\hat{f}$ when f is a step function of the form

$$f(t) = \begin{cases} c_j & t \in [a_{j-1}, a_j) \ (j=2 \ldots n) \\ 0 & t \notin [a_1, a_n) \end{cases}$$

where $0 \leq a_1 < a_2 < \ldots < a_n < \infty$.

For such a function f take for $\hat{f}$

$$\hat{f}(\omega) = \int_0^\infty f(t)x(dt,\omega) = \sum_{j=2}^{n} c_j[x^-(a_{j-1},\omega) - x^-(a_j,\omega)]$$

or, for that matter, any function in $\mathcal{L}^2_\Omega$ equal almost everywhere to this $\hat{f}$.

It follows from the assumption that the process has orthogonal increments that

$$\int_\Omega \hat{f}(\omega)g(\omega)p(d\omega) = \int_0^\infty f(t)g(t)dt$$

and this mapping is an isometry and, hence continuous. Therefore this map of $\mathcal{L}^2_F$ into $\mathcal{L}^2_\Omega$ can be extended to the closure of the set of step functions of the above type (in the topology generated by the scalar product $\int fgdF$). That is, if $\lim_{n\to\infty} f = f$ in the $\mathcal{L}^2_F$ topology then $\underset{n\to\infty}{\text{L.i.m.}}\ \hat{f}_n$ exists and we define

$$\hat{f} = \underset{n\to\infty}{\text{L.i.m.}}\ \hat{f}_n \ .$$

As the closure of this set of step functions is all of $\mathcal{L}^2_F$ we have an isometric mapping of $\mathcal{L}^2_F$ into $\mathcal{L}^2_\Omega$. It is interesting to note that, if the process is such that

$$\int_{\Omega}[x(t,\omega) - x(s,\omega)]p(d\omega) = 0 ,$$

(for example if it is a Brownian Motion Process) then, for every $f \in \mathcal{L}_F^2$,

$$\int_{\Omega} f(\omega)p(d\omega) = \int_{\Omega}\left[\int_0^{\infty} f(t)x(dt,\omega)\right]p(d\omega) = 0 .$$

Chapter XII Integration Over Outer Hilbert Space

8 Introduction.

In earlier lectures a theory of integration over Hilbert space (with respect to a Gaussian measure) was developed, and certain basic results derived. Among these was the fact that a cylinder function f(x) which is "integrable" possesses very strong "invariance" properties. Thus, in particular, it followed that the measure of cylinders which are cylinders with reference to a given basis (or, rather, with reference to a complete basic system) may be computed equally well with reference to any other complete basic system. This intimates that in some sense one should have "total additivity of the measure" even though, as we have seen, the measure we have defined in Hilbert space is not totally additive. This intimation would be substantiated if we could find an "extension" of Hilbert space — call it W ($W \supset \mathfrak{H}$) — such that the cylinders of $\mathfrak{H}$ (i.e., cylinders with reference to any base) are filled out in W by the addition of new elements (if C is a cylinder in $\mathfrak{H}$, call C_W the corresponding cylinder in W), in such a way that there would be a totally additive measure μ_W over W with respect to which the cylinders C_W would be measurable and such that

$$(*) \qquad \mu_W(C_W) = \mu_{\mathfrak{H}}(C) \ .$$

This goal will be, in part, achieved in the ensuing discussion.

For a given complete basic system (P), an extension W(P) of $\mathfrak{H}$ is constructed by considering so-called "weak elements". This space W(P), which we shall call a "basic corona", is simply an unrestricted sequence space in which the Hilbert space $\mathfrak{H}$ is imbedded according to its coordinatization relative to the basis corresponding to (P). The extension of the measure $\mu_{\mathfrak{H}}$ from $\mathfrak{H}$ to W(P) can be performed in a well-known manner so that it is totally additive and satisfies condition (*).

Next we consider simultaneously two complete basic systems, (P) and (Q), and the corresponding extensions, W(P) and W(Q). We then proceed to show that there exists a 1-1 measure-preserving map between them, called "natural transfer", which is defined "almost everywhere". Those elements in W(P) which admit natural transfer to W(Q) will then be identified with their images in W(Q); such elements will be called "common weak elements". The space W(P)(Q) of these common weak elements will be called the "corona generated by (P) and (Q)". The extension of the measure $\mu_{\mathfrak{h}}$ to W(P), when restricted to W(P)(Q), then leads to an extension of $\mu_{\mathfrak{h}}$ to W(P)(Q) which is totally additive and satisfies condition (*).

Finally, to any countable set $\{(P),(Q),\ldots\}$ of basic systems, we assign a corona W = W(P)(Q)... to which $\mu_{\mathfrak{h}}$ may be extended in such a way as to be totally additive and satisfy condition (*).

The relationship between integration over Hilbert space and integration over a corona will be investigated.

§2. Basic Extensions.

We first consider the case of constructing the corona of $\mathfrak{h}$ compatible with a single complete basic system $(P) = \{P^\sigma\}$. We shall refer to this as a basic corona or (P)-corona.

Definition of a "weak element". A weak element X relative to (P) is given by a sequence $\{x^\tau\}$ of elements of $\mathfrak{h}$ such that

$$P^\sigma x^\tau = x^\sigma \quad , \quad \text{for all } \sigma,\tau \text{ with } \sigma \leq \tau. \tag{2.1}$$

From the above definition it follows that each x^σ is in the subspace $\mathfrak{P}^\sigma = P^\sigma\mathfrak{h}$ and, setting

$$P^\sigma X = x^\sigma \ ,$$

we may associate with a weak element X a sequence $(x_1,x_2,\ldots)$ such that

$$P^{\sigma}X = x^{\sigma} = \sum_{i=1}^{\sigma} x_i p_i$$

(where $\{p_1, p_2, \ldots\}$ is the fixed basis associated with (P)). Also, in this way any sequence $(x_1, x_2, \ldots)$ determines a unique weak element. Clearly, for every $P \in (P)$ and every weak element X, $PX \in \mathfrak{h}$.

W(P) is the set of all weak elements relative to (P). Thus $W(P) \supset \mathfrak{h}$, and we see that we may view W(P) as the unrestricted sequence space and $\mathfrak{h}$ as the ℓ_2 subspace wherein $\{x_1, x_2, \ldots\}$ are the coordinates relative to the basis $\{p_1, p_2, \ldots\}$.

Definition: A cylinder in W(P) is a set whose characteristic function $\chi(X)$ satisfies $\chi(PX) = \chi(X)$ for some P in (P) and for all X in W(P).

For any such cylinder $C_{(P)} = C_{W(P)}$ in $W = W(P)$ we define the measure

$$\mu_{(P)}\left(C_{(P)}\right) = \mu_{W(P)}\left(C_{(P)}\right) = \mu_{\mathfrak{h}}\left(PC_{W(P)}\right)$$

where the P operating on $C_{W(P)}$ in the argument on the far right is that P in (P) such that $\chi(PX) = \chi(X)$. Hence, for the "projection", $C_{\mathfrak{h}} = C_{(P)} \cap \mathfrak{h}$, of $C_{(P)}$ into $\mathfrak{h}$, we have

$$\mu_{\mathfrak{h}}\left(C_{\mathfrak{h}}\right) = \mu_{\mathfrak{h}}\left(PC_{(P)}\right) = \mu_{(P)}\left(C_{(P)}\right) .$$

Since, via Kolmogorov's theorem (p. 11, Foundations of the Theory of Probability), these given measures over the cylinders $C_{(P)}$ can be extended to a countably additive measure we see that W(P) is indeed the desired basic corona.

§3. Transfer mappings.

The decisive step in our ultimate argument will be the construction of a corona of $\mathfrak{h}$ compatible with two given complete basic systems (P) and (Q). This construction will, essentially, reduct to showing that there exists a "transference mapping" between W(P) and W(Q) which leaves $\mathfrak{h}$ fixed. In this section we set up the machinery necessary for treating this transference.

Definition: Let (P) be a given complete basic system, W(P) the corresponding basic corona, X ε W(P), Q any finite dimensional projection of $\mathcal{H}$ and $\{\sigma_\nu\}$ a given subsequence of the integers. We say that "Q is applicable to X relative to $\{\sigma_\nu\}$" if and only if

$$\lim_{\nu\to\infty} QP^{\sigma_\nu} X$$

exists, in which case we write QX for this limit and write

$$QP^{\sigma_\nu} X \to QX .$$

Lemma 3.1: Given a complete basic system (P), a projection Q, and any sequence $\{\sigma_\nu\}$, let S_Q denote the set of all X ε W(P) such that Q is applicable to X relative to $\{\sigma_\nu\}$; then the difference, W(P) - S_Q, has W(P)-measure zero so that

$$\mu_{(P)}(S_Q) = 1 .$$

Proof: It clearly suffices to prove the lemma in the case where $\{\sigma_\nu\}$ is the sequence of all integers. Let $\{p_1, p_2, \ldots\}$ be the orthonormal basis of $\mathcal{H}$ corresponding to (P); and $\{q_1, \ldots, q_n\}$ an orthonormal basis for the range of Q. Then

$$QP^\sigma X = \sum_{i=1}^{n} \left(P^\sigma X, q_i\ q_i\right) = \sum_{i=1}^{n} q_i \sum_{j=1}^{\sigma} x_j h_{ij}$$

where $h_{ij} = (q_i, p_j)$. Thus the existence of $\lim_{\sigma\to\infty} QP^\sigma X$ is equivalent to the convergence, for each i = 1,...,n, of the infinite series

$$\sum_{j=1}^{\infty} x_j h_{ij} .$$

To prove the lemma one need only establish that each of these series converges for a set of $\mu_{(P)}$-measure 1. However, this is provided by the Three Series Theorem (for which see page 199 of Halmos, Measure Theory) since $\sum_j h_{ij}^2 = 1$.

Definition: Given a sequence $\{\sigma_\nu\}$, and two complete basic systems (P) and (Q). Let $X \in W(P)$ be such that, for all τ, Q^τ is applicable to X relative to $\{\sigma_\nu\}$. Then we say that "X is transferable to W(Q)" relative to $\{\sigma_\nu\}$.

Lemma 3.2: If $X \in W(P)$ is transferable to W(Q) relative to $\{\sigma_\nu\}$, then the sequence $\{Q^\tau X\}$ determines an element of W(Q). (We shall call the element $Y = QX$ the "transfer of X to W(Q) relative to $\{\sigma_\nu\}$".)

Proof: We need only observe that, for $\tau \geq \tau'$,

$$Q^{\tau'} Q^\tau X = Q^{\tau'} X$$

as

$$Q^{\tau'} Q^\tau X = Q^{\tau'} \lim_{\nu\to\infty} Q^\tau P^{\sigma_\nu} X$$

$$= \lim_{\nu\to\infty} Q^{\tau'} \left(Q^\tau P^{\sigma_\nu} X\right)$$

$$= \lim_{\nu\to\infty} Q^{\tau'} P^{\sigma_\nu} X = Q^{\tau'} X .$$

At this point we note that every $X \in W(P)$ which is is transferable to W(Q) and the transfer of X to W(Q) is X itself (given, now, in the (Q) coordinate system).

Lemma 3.3: Let $\mathcal{T} = \mathcal{T}(P)(Q)$ be the set of elements of W(P) which are transferable to W(Q) relative to some (any) given sequence $\{\sigma_\nu\}$; then

$$\mu_{(P)}(\mathcal{T}) = 1 .$$

Proof: The set $\mathcal{T}$ is evidently the intersection of all the sets S_{Q^τ} defined as in Lemma 3.1. Since the measure is totally additive in W(P) the statement of Lemma 3.3 follows from Lemma 3.1.

Remark: The statement of Lemma 3.3 applies, in particular, to transfer with respect to the sequence of all integers.

Definition: If $X \in W(P)$ is transferable to W(Q) relative to $\{\sigma_\nu\}$, and Y is its transfer we write

$$X \underset{\{\sigma_\nu\}}{\sim} Y$$

and call this mapping the "transfer relative to $\{\sigma_\nu\}$" from W(P) to W(Q).

In terms of the above notion, Lemma 3.3 asserts that the transfer of W(P) to W(Q) relative to $\{\sigma_\nu\}$ is defined almost everywhere on W(P), (that is, almost all of W(P) is transferred). Note again that any "transfer" leaves each element of $\mathfrak{h}$ fixed.

§4. Symmetric transfer.

A natural question which arises concerning transfers is whether or not they are 1-1, possibly apart from sets of measure 0. This is, in fact, true for transfer not only relative to some sequence $\{\sigma_\nu\}$ but even relative to the particular sequence consisting of all integers. As a first step in the proof of this we shall prove in this section that there exists a subsequence $\{\sigma_\nu\}$ (depending on the (P) and (Q) involved) relative to which this is true.

Theorem 4.1: Given any two complete basic systems, (P) and (Q), there exists a subsequence $\{\sigma_\nu\}$ of the integers such that, writing $\mathcal{A}(P)$ for the set of all $X \in W(P)$ for which there exists a $Y \in W(Q)$ such that

$$X \underset{\{\sigma_\nu\}}{\sim} Y \quad \text{and} \quad Y \underset{\{\sigma_\nu\}}{\sim} X$$

and writing (Q) for the set of all $Y \in W(Q)$ for which there exists an $X \in W(P)$ such that

$$Y \underset{\{\sigma_\nu\}}{\sim} X \quad \text{and} \quad X \underset{\{\sigma_\nu\}}{\sim} Y ,$$

one has

$$\mu_{(P)}(\mathcal{A}(P)) = \mu_{(Q)}(\mathcal{A}(Q)) = 1 .$$

In other words, given any two complete basic systems, (P) and (Q), there exists a subsequence $\{\sigma_\nu\}$ of the integers relative to which the transfer from W(P) to W(Q) is possible and is the inverse of that from W(Q) to W(P), apart from sets of measure 0. Thus, apart from sets of measure 0, these transfers are 1-1. Note also that, since the elements of $\mathfrak{f}$ are transferred into themselves here, as in our future discussions, all the sets of measure 0 which we exclude contain only weak elements outside $\mathfrak{f}$.

Proof of Theorem: Let $\{q_i\}$ be the basis corresponding to (Q) and $\{p_i\}$ the basis corresponding to (P). It suffices to prove that, for every fixed integer $\lambda \geq 1$, we can extract from any given sequence a subsequence $\{\tau_\nu\}$ such that

$$\mu_{(P)}\left\{X\epsilon W(P):\ \lim_{\nu\to\infty}\ \lim_{\tau\to\infty}\ P^\lambda Q^{\tau_\nu} P^\sigma X = P^\lambda X\right\} = 1 .$$

We have, for every τ and σ,

$$Q^\tau P^\sigma X = \sum_{i=1}^{\tau} (P^\sigma X, q_i) q_i$$

$$= \sum_{i=1}^{\tau} q_i \sum_{j=1}^{\sigma} x_j h_{ij} ,$$

where, as before, $h_{ij} = (q_i, p_j)$ so that, for each λ,

$$P^\lambda Q^\tau P^\sigma X = \sum_{i=1}^{\tau} \sum_{k=1}^{\lambda} h_{ik} p_k \sum_{j=1}^{\sigma} x_j h_{ij}$$

$$= \sum_{k=1}^{\lambda} p_k \sum_{j=1}^{\sigma} x_j \left[\sum_{i=1}^{\tau} h_{ik} h_{ij}\right]$$

For a set of X with $\mu_{(P)}$-measure 1, according to Lemma 3.1, the expression

$$\sum_{k=1}^{\lambda} p_k \left\{\sum_{j=1}^{\infty} x_j \left[\sum_{i=1}^{\tau} h_{ik} h_{ij}\right]\right\}$$

exists for each τ and λ. Thus, it suffices to show that, for each fixed k, we can find, out of any given sequence, a subsequence $\{\tau_\nu\}$, such that

$$\lim_{\nu\to\infty} \sum_{j=1}^{\infty} x_j \left[\sum_{i=1}^{\tau_\nu} h_{ik}h_{ij}\right] = x_k \tag{4.1}$$

occurs with $\mu_{(P)}$-measure 1. For, in a well-known manner, we can then find a common subsequence for which the relation (4.1) holds for all k.

For fixed k, we write

$$g_j(\tau) = \sum_{i=\tau+1}^{\infty} h_{ik}h_{ij}$$

so that (4.1) may be restated as

$$\lim_{\nu\to\infty} \sum_{j=1}^{\infty} x_j g_j(\tau_\nu) = 0 \tag{4.2}$$

By applying the unitary transformation $U = (h_{ij})$ to the vector $(\underbrace{0,\ldots,0}_{\tau},h_{\tau+1,k},\ldots)$ and noting that this is an isometry, we obtain

$$\sum_{j=1}^{\infty} g_j^2(\tau) = \sum_{i=\tau+1}^{\infty} h_{ik}^2 < \infty \ .$$

Thus it is clear that the distribution function of the random variable, $F(X) = \sum_{j=1}^{\infty} x_j g_j(\tau)$, is a normal distribution with mean 0 and variance $\sum_{j=1}^{\infty} g_j^2(\tau)$.

Applying Chebyshev's inequality (for which see page 200 of Halmos, Measure Theory) we observe that the inequality

$$\mu_{(P)}\left\{X: \left|\sum_{j=1}^{\infty} x_j g_j(\tau)\right| > \varepsilon\right\} \le \varepsilon^{-2} \sum_{i=\tau+1}^{\infty} h_{ik}^2$$

holds for any $\varepsilon > 0$ and that the right side of the inequality tends to 0 as $\tau \to \infty$. Therefore, according to a well known

theorem of probability theory (again see Halmos, Measure Theory) we can extract from any sequence a subsequence $\{\tau_\nu\}$ such that (4.2) holds with $\mu_{(P)}$-measure 1.

The above, together with suitable diagonalization arguments (see above) in connection with choices of subsequences, then leads to the desired result.

§5. Symmetric transfer is measure preserving.

We next wish to show that if the sequence $\{\sigma_\nu\}$ is chosen as provided by Theorem 4.1, the 1-1 map of the corresponding transfers (apart from sets of measure 0) between $W(P)$ and $W(Q)$ is measure preserving. For this purpose it suffices to prove

Theorem 5.1: For any $\mu_{(P)}$-Borel-measurable cylinder $C_{(P)} = C_{W(P)}$ of $W(P)$ its image $C_{(Q)}$ under this transfer is a $\mu_{(Q)}$-Borel-measurable set and

$$\mu_{(P)}(C_{(P)}) = \mu_{(Q)}(C_{(Q)}) \,. \tag{5.1}$$

The reason is that, from this last assertion together with the symmetry of the transfer, one can easily deduce that the exterior measure $\bar{\mu}_{(P)}$ of any subset of $W(P)$ equals the exterior measure $\bar{\mu}_{(Q)}$ of its transfer to $W(Q)$. The same statement then follows immediately for the inner measures and the transfer is measure preserving.

To prove the $\mu_{(Q)}$-measurability of the image of $C_{(P)}$ it clearly suffices to consider only cylinders $C_{(P)}$ defined by inequalities of the form

$$a_k < x_k < \beta_k \,, \quad k = 1,\ldots,n \quad (n \text{ fixed}) \,.$$

The image $C_{(Q)}$ is characterized by the fact that (apart from sets of measure 0) the re-transfers of its elements to $W(P)$ exist and lie in $C_{(P)}$. If $Y = (y_1, y_2, \ldots) \in C_{(Q)}$ is such an element, its transfer back to $W(P)$, (relative to $\{\sigma_\nu\}$, of course) has components

$$(5.2) \qquad \lim_{\nu\to\infty} \sum_{i=1}^{\sigma_\nu} y_i h_{ik} \; ; \qquad k = 1,\dots,n \; .$$

Thus, apart from sets of measure 0, $C_{(Q)}$ is determined by the conditions

$$(5.3) \qquad \alpha_k < \lim_{\nu\to\infty} \sum_{i=1}^{\sigma_\nu} y_i h_{ik} < \beta_k \; ; \quad k = 1,\dots,n \; .$$

That $C_{(Q)}$ is $\mu_{(Q)}$-Borel-measurable then follows immediately from the fact that

$$(5.4) \quad C_{(Q)} = \bigcup_{m=1}^{\infty} \bigcup_{\sigma_\nu \geq m} \left\{ Y: \; \alpha_k + \frac{1}{m} < \sum_{i=1}^{\sigma_\nu} y_i h_{ik} < \beta_k - \frac{1}{m} \; ; \; k=1,\dots,n \right\}$$

where the set, of $\mu_{(Q)}$-measure 0, elements Y for which any of the $\lim_{\nu\to\infty} \sum_{i=1}^{\sigma_\nu} y_i h_{ik}$ do not exist is ignored. We shall systematically ignore such sets of measure 0, since any errors commited on their behalf does not disturb the validity of our proofs.

The proof of (5.1) will now proceed with the aid of two lemmas.

Lemma 5.1: Let

$$A_{\sigma_\nu} = A_{\sigma_\nu}[\alpha_1,\beta_1,\dots,\alpha_n,\beta_n]$$

be the set

$$A_{\sigma_\nu} = \left\{ Y: \; \alpha_k < \sum_{i=1}^{\sigma_\nu} y_i h_{ik} < \beta_k \; ; \quad k = 1,\dots,n \right\} \; ;$$

then

$$(5.5) \qquad \lim_{\nu\to\infty} \mu_{(Q)}(A_{\sigma_\nu}) = \mu_{(P)}(C_P) \; .$$

Proof: Let $\tilde{A}_{\sigma_\nu} = A_{\sigma_\nu} \cap \mathfrak{H} = \left\{ X \in \mathfrak{H}: \; P\, Q^{\sigma_\nu} X \in I \right\}$

and let

$$\tilde{C}_{(P)} = C_{(P)} \cap \mathfrak{H} = \left\{ X \in \mathfrak{H}: \; P^{\sigma} X \in I \right\}$$

where I is the quasi-interval in $P^{\sigma} = P^{\sigma}\hbar$ such that the coordinates (relative to the fixed basis of P^{σ} implied here) are in the intervals (α_k, β_k), $k = 1,\ldots,n$.

Thus, if $g(x)$ is the characteristic function of $\tilde{C}_{(P)}$, it is a cylinder function over $\hbar$, such that

$$g(P^{\sigma}x) = g(x)$$

Hence,

$$g\left(Q^{\sigma_\nu} x\right) = g\left(P^{n}Q^{\sigma_\nu} x\right)$$

is the characteristic function of $\tilde{A}_{\sigma_\nu}$.

Using the fact that cylinder functions which are "integrable" (and $g(x)$ is integrable) are invariant, we get, by Theorem A,

$$\mu_{(P)}(C_{(P)}) = \mu_{\hbar}(\tilde{C}_{(P)}) = \int_{P^{\sigma}} g(x)\, d\mu_{P^{\sigma}}$$

$$= \lim_{\nu\to\infty} \int_{Q^{\sigma_\nu}} g(x)\, d\mu_{Q^{\sigma_\nu}}$$

$$= \lim_{\nu\to\infty} \mu_{\hbar}(\tilde{A}_{\sigma_\nu}) = \lim_{\nu\to\infty} \mu_{(Q)}(A_{\sigma_\nu})$$

Lemma 5.2: Let

$$E = \left\{ Y: \begin{array}{l} \alpha_k \le \lim\limits_{\nu\to\infty} \sum\limits_{i=1}^{\sigma_\nu} y_{ik} \le \beta_k\,, \quad k = 1,\ldots,n\,; \\ \text{and } \underline{\text{at least one}} \text{ of the equality signs holds} \end{array} \right\};$$

then

$$\mu_{(Q)}(E) = 0\,.$$

Proof: It suffices to prove that the set,

$$E_1 = \left\{ Y: \begin{array}{l} \lim\limits_{\nu\to\infty} \sum\limits_{i=1}^{\sigma_\nu} y_i h_{i1} = c \ ; \\ \\ a_k \le \lim\limits_{\nu\to\infty} \sum\limits_{i=1}^{\sigma_\nu} y_i h_{ik} \le \beta_k \ ; \quad k = 2,\ldots,n \end{array} \right\} ,$$

has $\mu_{(Q)}$-measure 0. For any $\varepsilon > 0$, take $\hat{a}_1 = c-\varepsilon$, $\hat{\beta}_1 = c+\varepsilon$; $\hat{a}_k = a_k-\varepsilon$, $\hat{\beta}_k = \beta_k+\varepsilon$, $k=2,\ldots,n$. Then, setting $\hat{A}_{\sigma_\nu} = A_{\sigma_\nu}[\hat{a}_1,\hat{\beta}_1,\ldots,\hat{a}_n,\hat{\beta}_n]$, we see that

$$E_1 \subset \varlimsup_{\nu\to\infty} \hat{A}_{\sigma_\nu}$$

and, hence,

$$\mu_{(Q)}(E_1) \le \mu_{(Q)} \varlimsup_{\nu\to\infty} \left(\hat{A}_{\sigma_\nu}\right) \le \varlimsup_{\nu\to\infty} \mu_{(Q)}\left(\hat{A}_{\sigma_\nu}\right) .$$

Applying Lemma 5.1, this in turn yields

$$\mu_{(Q)}(E_1) \le \mu_{(P)}\left(\hat{C}_{(P)}\right)$$

where $\hat{C}_{(P)} = \left\{Y: \ \hat{a}_k < x_k < \hat{\beta}_k; \ k = 1,\ldots,n\right\}$. Finally, since $\hat{\beta}_1-\hat{a}_1 = 2\varepsilon$, we note that $\lim\limits_{\varepsilon\to 0} \mu_{(P)}(\hat{C}_{(P)}) = 0$; and hence $\mu_{(Q)}(E_1) = 0$. Thus, Lemma 5.2 is proved.

We may now obtain (5.1) rather simply. From

$$C_{(Q)} \subset \left[\varlimsup_{\nu\to\infty} A_{\sigma_\nu}\right] \subset \left[C_{(Q)} \cup E\right]$$

it follows that

$$\mu_{(Q)}(C_{(Q)}) = \mu_{(Q)}\left(\varlimsup_{\nu\to\infty} A_{\sigma_\nu}\right) . \tag{5.6}$$

Also

$$\left[\varlimsup_{\nu\to\infty} A_{\sigma_\nu}\right] \subset \left(\left[\varlimsup_{\nu\to\infty} A_{\sigma_\nu}\right]\right) \cup E \ ,$$

so that, using Lemma 5.2,

$$\mu_{(Q)}\left(\varlimsup_{\nu\to\infty} A_{\sigma_\nu}\right) = \mu_{(Q)}\left(\varlimsup_{\nu\to\infty} A_{\sigma_\nu}\right) . \tag{5.7}$$

From (5.6), (5.7), and Lemma 5.1 we conclude that

$$\mu_{(Q)}(C_{(Q)}) = \mu_{(Q)}\left(\varliminf_{\nu\to\infty} A_{\sigma_\nu}\right) = \mu_{(Q)}\left(\overline{\lim_{\nu\to\infty}} A_{\sigma_\nu}\right)$$
$$= \lim_{\nu\to\infty} \mu_{(Q)}\left(A_{\sigma_\nu}\right) = \mu_{(P)}(C_{(P)}) .$$

Thus Theorem 5.1 is proved.

§6. Natural transfer.

As was mentioned in §4, transfer relative to the sequence of all integers is symmetric; that is, the use of the subsequence $\{\sigma_\nu\}$ may be eliminated. This may now be achieved rather easily on the basis of the result of §5. This transfer relative to all integers will be referred to as "natural transfer". We shall denote natural transfer by $X \rightsquigarrow Y$.

Denote by S' the set of those $X \in W(P)$ which are transferable to W(Q) relative to the sequence $\{\sigma_\nu\}$ (used for symmetric transfer); and by S the set of $X \in W(P)$ transferable to W(Q) relative to the sequence of all integers. Clearly $S' \supset S$, and hence, from the remark made following Lemma 3.3, we infer that

$$\mu_{(P)}(S') = \mu_{(P)}(S) = 1$$

or

$$\mu_{(P)}(S'-S) = 0 .$$

Let $\widehat{S'-S}$ denote the transfer to W(Q), relative to $\{\sigma_\nu\}$, of (S'-S). Since this transfer preserves measure, we obtain

$$\mu_{(Q)}(\widehat{S'-S}) = 0 .$$

Thus we see that the extra images provided by transferring relative to the sequence $\{\sigma_\nu\}$ (instead of all integers) have $\mu_{(Q)}$-measure 0, and similarly for transfer from W(Q) to W(P). From this we conclude that symmetric transfer may be made relative to all integers with measure 1.

Theorem 6.1: Given any two complete basic systems, (P) and (Q), if

$$\hat{\mathcal{A}}(P) = \left\{ X \epsilon W(P): \begin{array}{l} \text{there exists a } Y \epsilon W(Q) \text{ such that} \\ X \rightsquigarrow Y \text{ and } Y \rightsquigarrow X \end{array} \right\}$$

and

$$\hat{\mathcal{A}}(Q) = \left\{ Y \epsilon W(Q): \begin{array}{l} \text{there exists an } X \epsilon W(P) \text{ such that} \\ Y \rightsquigarrow X \text{ and } X \rightsquigarrow Y \end{array} \right\} ;$$

then

$$\mu_{(P)}(\mathcal{A}(P)) = \mu_{(Q)}(\mathcal{A}(Q)) = 1 .$$

§7. Preservation of Weak Convergence under Natural Transfer.

Definition: If $\{X^{(n)}\}$ is a sequence of points of W(P), then we say that $X^{(n)} \rightarrow X \epsilon W(P)$ weakly (weak convergence) if and only if, for every k, $P^k X^{(n)} \rightarrow P^k X$ in the Hilbert space norm.

If $\{X^{(n)}\}$ is a sequence of elements of W(P) such that $X^{(n)} \rightarrow X$ weakly; and such that under natural transfer

$$\text{(a)} \qquad X^{(n)} \rightsquigarrow Y^{(n)} \epsilon W(Q)$$

and

$$\text{(b)} \qquad X \rightsquigarrow Y \epsilon W(Q) ,$$

it is natural to ask whether or not the sequence $\{Y^{(n)}\}$ converges weakly in W(Q), and, if it does, whether or not the limit is Y. We shall show that in "a certain sense" these questions may be answered in the affirmative, and that these properties characterize the natural transfer.

Theorem 7.1: Given any $\epsilon > 0$, there exists an "outside" set, S_ϵ of elements of W(P), with $\mu_{(P)}(S_\epsilon) < \epsilon$, such that, if

$$(7.1) \quad X^{(n)} \rightarrow X \text{ weakly, } \left[X^{(n)}, X \epsilon W(P)\right] ,$$

$$X^{(n)} \rightsquigarrow Y^{(n)}, \quad \left[Y^{(n)} \epsilon W(Q)\right] ,$$

and

(7.2) $X^{(n)}, X \notin S_\varepsilon$,

then

(7.3) $Y^{(n)} \to Y$ weakly.

We shall say simply that natural transfers preserve weak convergence outside of S_ε.

Proof: We exclude to begin with the set of $\mu_{(P)}$-measure 0 on which the natural transfer between $W(P)$ and $W(Q)$ is not symmetric. Assuming (7.1) and (7.2), we recall first that

$$Q^k Y^{(n)} = \sum_{i=1}^{k} \left[\sum_{j=1}^{\infty} x_j^{(n)} h_{ij} \right] q_i$$

and

$$Q^k Y = \sum_{i=1}^{k} \left[\sum_{j=1}^{\infty} x_j h_{ij} \right] q_i$$

so that to derive (7.3) it suffices to show that, for each fixed k,

$$c_k^{(n)} = \sum_{j=1}^{\infty} x_j^{(n)} h_{kj} \text{ converges, as } n \to \infty, \text{ to } c_k = \sum_{j=1}^{\infty} x_j h_{kj}$$

Since $\sum_{j=1}^{\infty} x_j^{(n)} h_{kj}$ converges with $\mu_{(P)}$-measure 1, given any $\varepsilon > 0$ and $\delta > 0$ there exists an integer M and a set $S = S_{\varepsilon,\delta,k}$, with $\mu_{(P)}(S) < \varepsilon/2^{(k+1)/\delta}$, such that, for all $X^{(n)} \notin S$,

$$(7.4) \qquad \left| \sum_{j=1}^{N} x_j^{(n)} h_{kj} - c_k^{(n)} \right| < \delta \quad \text{for all } N \geq M .$$

Fixing the N, since $X^{(n)} \to X$ weakly, it follows that, for all $N \geq M$ and all $n \geq n' = n'(N)$,

$$(7.4) \qquad \left| \sum_{j=1}^{N} x_j h_{kj} - c_k^{(n)} \right| < 2\delta .$$

Since we have assumed $X \prec Y$, $c_k = \sum_{j=1}^{\infty} x_j h_{kj}$ converges, and hence there exists an $N' \geq M$ such that

$$(7.5) \qquad \left| \sum_{j=1}^{N'} x_j h_{kj} - c_k \right| < \delta .$$

Then (7.4) and (7.5) yield

$$(7.6) \qquad |c_k^{(n)} - c_k| < 3\delta \quad \text{for all } n \geq n'(N') .$$

Thus if we choose successively, $\delta = 1, \frac{1}{2}, \frac{1}{3}, \ldots$ and exclude the $X^{(n)}$ successively from sets of measure $< \frac{\varepsilon}{2^{k+1}}, \frac{\varepsilon}{2^{k+2}}, \frac{\varepsilon}{2^{k+3}}, \ldots$ we exclude in all a set of measure $< \varepsilon/2^k$, and from (7.6) we get that $c_k^{(n)} \to c_k$. Since we must do this for each k we exclude altogether a set of measure $< \sum_k \varepsilon/2^k = \varepsilon$. A simple argument shows, also, that no elements of $\mathfrak{h}$ need be excluded.

Theorem 7.2: Apart from sets of measure 0, the natural transfer is the only mapping of W(P) into W(Q) defined in W(P) except for an outside set of $\mu_{(P)}$-measure 0, which leaves $\mathfrak{h}$ fixed, and preserves weak convergence in the sense described in Theorem 7.1.

Proof: Given an $X \in W(P)$ which is transferred to W(Q) under the given mapping, we have

$$X^{(n)} = P^{(n)}X \to X \quad \text{weakly} .$$

If $Y \in W(Q)$ is the image of X under the transfer, and $Y^{(n)}$ is the image of $X^{(n)}$, then, since $\mathfrak{h}$ is left fixed, and $X^{(n)} \in \mathfrak{h}$,

$$Y^{(n)} = \sum_{j=1}^{m} x_j p_j = \sum_{i=1}^{\infty} q_i \sum_{j=1}^{m} x_j h_{ij}$$

Then by assumption $Y^{(n)} \to Y$ weakly and we see that the components of Y are $\sum_{j=1}^{\infty} x_j h_{ij}$, which is precisely what they would be under natural transfer.

§8. Transitivity of Natural Transfer.

Suppose that we have three complete basic systems (P), (Q) and (R). Then a natural transfer will allow us to pass from W(P) to W(Q) and from W(Q) to W(R); or

(8.1) $\quad X \sim Y \sim Z\ ,\quad X \in W(P)\ ,\quad Y \in W(Q)\ ,\quad Z \in W(R)\ .$

This determines a map $X \sim Z$ from W(P) to W(R). That this map is the natural transfer from W(P) to W(R) follows, since, clearly, the map is 1-1 apart from outside sets of measure zero, leaves $\mathfrak{h}$ fixed, and preserves weak convergence in the sense of Theorem 7.1.

§9. Construction of the corona.

We are now in a position to construct the extension of the Hilbert space $\mathfrak{h}$ which we have called a corona, compatible with a given sequence of complete basic systems $\{(P)_\nu\}$. From previous results we infer that, after discarding a countable number of outside sets of measure 0 in each local corona $W(P)_N$, the natural transfers between the $W(P)_N$ determine a symmetric and transitive relation between their elements. Thus we are led to form classes

$$\underline{X} = \left\{X^{(\nu)} \in W(P)_\nu\colon \text{ the } X^{(\nu)} \text{ transfer into one another}\right\}.$$

These clearly form a linear space W containing $\mathfrak{h}$ (each element of $\mathfrak{h}$ may be identified with such a class since natural transfers leave $\mathfrak{h}$ fixed).

Define a cylinder C_W of W to be a set of elements of W such that, for some ν, their $(P)_\nu$ representatives form a cylinder (except, possibly, for a set of $\mu_{(P)_\nu}$-measure 0) in $W(P)_\nu$. Then, assigning the measures

$$\mu_W(C_W) = \mu_{(P)_\nu}(C_{(P)_\nu})$$

to cylinders, we see that these generate a countably additive measure over W are required for the corona. We shall write symbolically

$$W = \bigcap_{\nu} W(P)_{\nu} .$$

Let us review, briefly, what has been done so far. We have shown that, given a Hilbert space $\mathfrak{h}$ with a unitarily invariant measure μ defined on it and a denumerable set of basic systems $\{(P)_{\nu}\}$ there is a linear space W and a measure μ_W defined on it such that

1) μ_W is totally additive,

2) $\mathfrak{h} \subset W$,

3) Every projection Q in each of the basic systems $(P)_{\nu}$ can be extended to a projection $\hat{Q}$ defined on W with a finite dimensional range $\hat{Q}W = Q\mathfrak{h} \subset \mathfrak{h}$,

4) There is a topology on W such that each such $\hat{Q}$ is continuous and which reduces, on $\mathfrak{h}$, to the regular Hilbert space topology,

5) For every Q as above and every measurable set $C \subset Q\mathfrak{h}$

$$\mu_W\{\underline{X} \in W: \hat{Q}\underline{X} \in C\} = \mu_{\mathfrak{h}}\{x \in \mathfrak{h}: Qx \in C\} ,$$

and

6) If, for any ν and ν', $\{p_k\}$ is the basis of $\mathfrak{h}$ corresponding to $(P)_{\nu}$, $\{p'_k\}$ that corresponding to $(P)_{\nu'}$, and U is a linear transformation such that $Up_k = p'_k$, then U may be extended linearly to a transformation $\hat{U}$ defined on W and, except for a set of μ_W-measure 0, $\hat{U}$ will be continuous in the topology of 4) and, for any measurable set $S \subset W$,

$$\mu_W(S) = \mu_W\{\underline{X} \in W: \hat{U}\underline{X} \in S\} .$$

Since the change described in 6) from one basic system to another is just a unitary transformation the situation may be equivalently described as follows. Given a Hilbert space $\mathfrak{h}$ with a

unitarily invariant measure $\mu_{\mathfrak{h}}$ and a denumerable group of unitary transformations, there exists an extension W of $\mathfrak{h}$ with a totally additive measure μ_W which is invariant under the linear extension to W of the group of transformations and, further, these extended transformations are continuous in the topology induced on W by projection into $\mathfrak{h}$.

§10. Extension of Integration over Hilbert space to Integration over the Corona.

We next wish to consider the question of replacing the integration of a function over the Hilbert space by the integration over a corona of a suitable extension of the function to the corona. For convenience we shall restrict our attention to the $\mathcal{L}_1$ theory.

We first consider the question with respect to a single basic system.

Theorem 10.1: Given a complete basic system (P) and a function $f(x)$ over $\mathfrak{h}$ such that $f \in \mathcal{L}_1^{(P)}$, then there exists a function, $\hat{f}(X) = \hat{f}(X|(P))$, defined over the basic corona W(P) such that

(1) $\hat{f}(P^\sigma x|(P)) = f(P^\sigma x)$ for all $x \in \mathfrak{h}$, and all σ,

(2) $\hat{f}(X)$ is integrable over W(P) with respect to $\mu_{(P)}$,

(3) $\hat{f}(P^\sigma X)$ converges in the $\mathcal{L}_1$ norm to $\hat{f}(X|(P))$,

and

(4) $\int_{\mathfrak{h}} f(x)d\mu_{(\mathfrak{h})} = \int_{W(P)} f(X)d\mu_{(P)}$.

Proof: For any X and σ, let

$$\hat{f}(P^\sigma X) = f(P^\sigma X) .$$

Then

$$\int_{W(P)} |\hat{f}(P^\sigma X)-\hat{f}(P^\tau X)|d\mu_{(P)} = \int_{\mathfrak{h}} |f(P^\sigma x)-f(P^\tau x)|d\mu_{\mathfrak{h}} .$$

The right side approaches 0 as σ and τ approach ∞ since $f \in \mathcal{L}_1^{(P)}$, so that the left side does also. However, since $\mu_{(P)}$

is a totally additive measure over $W(P)$, there exists a subsequence $\hat{f}(P^{\sigma_\nu} X)$ which converges almost everywhere to a function $\hat{f}(X)$ integrable over $W(P)$. Clearly this definition is consistent with that already given at the points $P^{\sigma}X$.

It then follows that

$$\lim_{\tau \to \infty} \int_{W(P)} |\hat{f}(X) - \hat{f}(P^{\tau}X)| d\mu_{(P)} = 0$$

so that

$$\int_{W(P)} \hat{f}(X) d\mu_{(P)} = \lim_{\tau \to \infty} \hat{f}(P^{\tau}X) d\mu_{(P)}$$

$$= \lim_{\tau \to \infty} \int_{\mathfrak{k}} f(P^{\tau}x) d\mu_{P^{\tau}}$$

$$= \int_{\mathfrak{k}} f(x) d\mu_{\mathfrak{k}} .$$

Remark: We see from this that the function $\hat{f}(X) = \hat{f}(X|(P))$ is uniquely determined almost everywhere by conditions 1), 2) and 3).

In the general case we have

Theorem 10.2: Let $(P)_\nu$ be a sequence of complete basic systems and let $W = \bigcap_\nu W(P)_\nu$ be the corresponding corona compatible with these. Let $f(x)$ be a function over $\mathfrak{k}$ which is invariant. For $\underline{X} \in W$ define $P_\nu^\sigma \underline{X} = P_\nu^\sigma X^{(\nu)}$ where $X^{(\nu)}$ is the component of $\underline{X}$ in $W(P)_\nu$. Then

(1) $\hat{f}(P_\nu^\sigma \underline{X}) = f(P_\nu^\sigma X^{(\nu)})$, for all ν, σ.

(2) $\hat{f}(\underline{X})$ is integrable over W with respect to μ_W.

(3) For each ν, the sequence $\hat{f}(P_\nu^\sigma(\underline{X})$ converges to $\mathcal{L}_1$ norm to $\hat{f}(\underline{X})$, and

(4) $\int_{\mathfrak{k}} f(x) d\mu_{\mathfrak{k}} = \int_W \hat{f}(\underline{X}) d\mu_W$.

Proof: The previous theorem provides functions $(X^{(\nu)}|(P)_\nu)$ defined in $W(P)_\nu$ such that

$$\int_{\Omega} f(x)d\mu_{\Omega} = \int_{W(P)_\nu} \hat{f}(X^{(\nu)}|(P)_\nu)d\mu_{(P)_\nu} .$$

The theorem then follows if we can show that if $X^{(\nu)} \hookleftarrow X^{(\tau)}$ under natural transfer then

$$\hat{f}(X^{(\nu)}|(P)_\nu) = \hat{f}(X^{(\tau)}|(P)_\tau)$$

almost everywhere with respect to $\mu_{(P)_\nu}$ (or $\mu_{(P)_\tau}$). Then, since

$$\lim_{\sigma\to\infty} \int_{W(P)_\nu} |\hat{f}(X^{(\nu)}|(P)_\nu) - f(P_\nu^\sigma X^{(\nu)})|d\mu_{(P)_\nu} = 0 ;$$

$$\lim_{\rho\to\infty} \int_{W(P)} |\hat{f}(X^{(\tau)}|(P)_\tau) - f(P_\tau^\rho X^{(\tau)}|d\mu_{(P)_\tau} = 0 ;$$

and from the invariance of f together with the fact that $X^{(\tau)}$ and $X^{(\nu)}$ are natural transfers of one another, we get that, in $W' = W(P)_\nu \cap W(P)_\tau$,

$$\lim_{\sigma,\rho\to\infty} \int_{W'} |f(P_\tau^\rho X) - f(P_\tau^\sigma X)|d\mu_{W'} = 0 .$$

Combining the above yields

$$\int_{W'} |\hat{f}(X^{(\nu)}|(P)_\nu) - \hat{f}(X^{(\tau)}|(P)_\tau|d\mu_{W'} = 0$$

and the desired result follows.

We have thus proved that for any extension W in the sense of §9 and any function f integrable over Ω there is an extension $\hat{f}$ of f (uniquely determined almost everywhere by conditions 1), 2), and 3) of Theorem 10.2) which is integrable over W with

$$\int_W f\, d\mu_W = \int_\Omega f\, d\mu_\Omega .$$

Chapter XIII

The Integrability of Polynomials*

In this chapter we shall prove, under suitable conditions on the coefficient functions, that, for each basic system (P) and for $\alpha = 2$, a "polynomial" or "product of polynomials" is integrable in the sense of Chapter V and that this integral is independent of (P) and that, further, it is invariant in the sense of V-6 relative to any system $\mathcal{R}$.

We recall from Chapter VII that we may treat polynomials in $\mathfrak{H}$ in terms of monomials of degree m. If we represent $\mathfrak{H}$ as $\mathcal{L}^2(\mathcal{S})$, the Hilbert space of square-integrable functions over the space $\mathcal{S}$, then these will be of the form

$$f[x] = \overbrace{\int_{\mathcal{S}} \cdots \int_{\mathcal{S}}}^{m} \phi(s_1, \ldots, s_{2m}) x(s_1) \ldots x(s_m) ds_1 \ldots ds_m \tag{13.1}$$

where

$$x = x(\cdot) \in \mathcal{L}^2(\mathcal{S})$$

represents an arbitrary element of $\mathfrak{H}$.

We wish to prove first the integrability of f under appropriate conditions on ϕ. We note that ϕ may be symmetrized in all its variables without altering f and hereafter we shall assume this has been done. Further, if $\phi \in \mathcal{L}^2(\mathcal{S}^m)$, the space of square-integrable functions of m variables each in $\mathcal{S}$, the integral in (13.1) will exist for all $x \in \mathcal{L}^2(\mathcal{S})$ and f is defined on all of $\mathfrak{H}$ and is independent of the particular representation.

For odd m it is clear from symmetry considerations that $\mathcal{J}(f) = 0$. We recall also [see VII-8] that for $m = 2n$, formal manipulation gives

---

* This chapter constitutes part of a thesis to be submitted to New York University in partial fulfillment of the requirements for the Ph.D. degree.

$$(13.2) \qquad \vartheta(f) = [(2n)!/2^n n!](\mathrm{Tr})^n \phi$$

where $(\mathrm{Tr})^\nu \phi$ is defined by

$$[\mathrm{Tr}\ \phi(s_1,\ldots,s_{2n})] = \phi_1(s_1,\ldots,s_{2n-2}) = \int_{\mathcal{A}} \phi(s_1,\ldots,s_{2n-2},\sigma,\sigma)d\sigma$$

$$(13.3) \qquad (\mathrm{Tr})^\nu \phi = \phi_\nu(s_1,\ldots,s_{2n-2\nu}) = \mathrm{Tr}\ \phi_{\nu-1} \ .$$

We must now show that for suitable ϕ the functional f is integrable and the integral, defined as

$$(13.4) \qquad \vartheta(f) = \lim_{k\to\infty} \vartheta(f[P^k \cdot]) \ ,$$

is correctly given by (13.2).

On finite dimensional spaces (13.2) holds, as may be verified by direct computation, the integrals in (13.3) reducing to finite sums. Specifically we note that if the range of P^k is spanned by the orthonormal set $\{p_1(\cdot),\ldots,p_k(\cdot)\}$ we may write

$$(13.5) \qquad f[P^k x] = \sum_{i_1=1}^{k} \cdots \sum_{i_{2n}=1}^{k} \phi^{(k)}_{i_1,\ldots,i_{2n}} \xi_{i_1} \cdots \xi_{i_{2n}}$$

for any $x \in \mathfrak{X}$ where

$$\xi_i = \int_{\mathcal{A}} x(s)p_i(s)ds \qquad \text{and}$$

$$(13.6) \qquad \phi^{(k)}_{i_1,\ldots,i_{2n}} = \int_{\mathcal{A}} \cdots \int_{\mathcal{A}} \phi(s_1,\ldots,s_{2n})p_{i_1}(s_1)\cdots p_{i_{2n}}(s_{2n})ds_1\cdots ds_{2n}$$

In this notation we have

$$(13.7) \qquad \vartheta(f[P^k \cdot]) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left\{ \sum_{i_1=1}^{k} \cdots \sum_{i_{2n}=1}^{k} \phi^{(k)}_{i_1,\ldots,i_{2n}} \xi_{i_1} \cdots \xi_{i_{2n}} \right\} \cdot$$

$$= \exp\left[-\frac{1}{2}\sum_{j=1}^{k} \xi_j^2\right] \sqrt{2\pi}^{-k}\, d\xi_1 \ldots d\xi_k$$

$$= \sum_{i_1=1}^{k} \cdots \sum_{i_n=1}^{k} \phi^{(k)}_{i_1 i_1 i_2 i_2 \cdots i_n i_n} = (\mathrm{Tr})^n \phi^{(k)}$$

making use of the symmetry of $\phi^{(k)}$ in its indices, which follows from that of ϕ.

In Chapter VII (pp. 3-6) conditions for the integrability of quadratic polynomials [n = 1] were established. These conditions may be stated in terms of linear operators associated with f. We shall similarly associate operators with polynomials of higher degree - though in this case the operator is defined over a Hilbert space other than $\mathfrak{H}$ - and prove a similar result.

We begin with some definitions.

Definition. To any function $\phi \in \mathcal{L}^2(\mathcal{I}^m)$ we associate linear operators L_ϕ^μ from $\mathcal{L}^2(\mathcal{I}^\mu)$ to $\mathcal{L}^2(\mathcal{I}^\nu)$, where $\mu \leq \nu = m-\mu$, by

$$(13.8)\quad L_\phi^\mu X^\mu = Y^\nu(t_1,\ldots,t_\nu) = \int_{\mathcal{I}}\cdots\int_{\mathcal{I}} \phi(t_1,\ldots,t_\nu;s_1,\ldots,s_\mu)X^\mu(s_1,\ldots,s_\mu)ds_1\ldots ds_\mu .$$

If we let $S^\mu = \{s_1,\ldots,s_\mu\}$ be a generic point of $\mathcal{I}^\mu$, etc., we may write $X^\mu(s_1,\ldots,s_\mu) = X^\mu(S^\mu)$, etc., and

$$(13.8')\qquad Y^\nu(T^\nu) = \int_{\mathcal{I}^\mu} \phi(T^\nu,S^\mu)X^\mu(S^\mu)dS^\mu$$

where $dS^\mu = ds_1\ldots ds_\mu$ and $\int_{\mathcal{I}^\mu} = \int_{\mathcal{I}}\cdots\int_{\mathcal{I}}$.

Definition. To any element $x = x(\cdot)$ of $\mathcal{L}^2(\mathcal{I})$ we associate elements X_x^μ of $\mathcal{L}^2(\mathcal{I}^\mu)$ for each μ by

$$(13.9)\qquad X_x^\mu(S^\mu) = x(s_1)x(s_2)\ldots x(s_\mu) .$$

That $X_x^\mu \in \mathcal{L}^2(\mathcal{I}^\mu)$ for $x \in \mathcal{L}^2(\mathcal{I})$ and any μ is clear.

We now note that, for m = 2n,

$$(13.10)\qquad f[x] = (X_x^n,\ L_\phi^n X_x^n)$$

and that

$$(13.11)\qquad (\mathrm{Tr})^n\phi = \mathrm{Tr}\ L_\phi^n$$

where the right hand side of (13.10) is the scalar product in $\mathcal{L}^2(\mathcal{S}^n)$ and the Tr on the right of (13.11) is the ordinary trace of a linear operator over $\mathcal{L}^2(\mathcal{S}^\mu)$ provided this exists. No confusion should arise from this dual use of the symbol (Tr).

Definition. To any basis $\{p_1,\ldots\}$ of $\mathfrak{H} = \mathcal{L}^2(\mathcal{S})$ we associate a basis $\{p_1^\mu,\ldots\}$ of $\mathcal{L}^2(\mathcal{S}^\mu)$ by indexing appropriately the set

$$\left\{p_\nu^\mu(S^\mu) = p_{i_1}(s_1)p_{i_2}(s_2)\ldots p_{i_\mu}(s_\mu) : i_1=1,2,\ldots,\ i_2=1,2,\ldots,\ldots,i_\mu=1,2,\ldots\right\}.$$

It is clear that this is a complete orthonormal basis of $\mathcal{L}^2(\mathcal{S}^\mu)$ if $\{p_1,\ldots\}$ is a complete orthonormal basis of $\mathfrak{H}$.

Definition. To any operator L on $\mathfrak{H}$ we associate, for each μ, an operator L_μ on $\mathcal{L}^2(\mathcal{S}^\mu)$ by

$$(13.12)\quad L_\mu X^\mu = L_\mu \sum_\nu a_\nu p_\nu^\mu = \sum_\nu a_\nu L_\mu(p_{i_1}p_{i_2}\ldots p_{i_\mu})$$

$$= \sum_\nu a_\nu(Lp_{i_1}Lp_{i_2}\ldots Lp_{i_\mu}) \quad .$$

It is clear that if L is linear on $\mathfrak{H}$ then L^μ is linear on $\mathcal{L}^2(\mathcal{S}^\mu)$.

We note that if X_1^μ, $X_2^\mu \in \mathcal{L}^2(\mathcal{S}^\mu)$ are of the form

$$(13.13)\qquad X_i^\mu(S^\mu) = x_i^1(s_1)x_i^2(s_2)\ldots x_i^\mu(s_\mu)$$

then

$$(13.14)\qquad (X_1^\mu, X_2^\mu) = (x_1^1, x_2^1)(x_1^2, x_2^2)\ldots(x_1^\mu, x_2^\mu) \quad .$$

We see that if P^k is an orthogonal projection with finite dimensional range so is P_μ^k and if U is a unitary operator

$$(13.15)\quad (U_\mu X^\mu, U_\mu Y^\mu) = (U_\mu \sum_\sigma a_\sigma p_\sigma^\mu, U_\mu \sum_\rho \beta_\rho p_\rho^\mu)$$

$$= \sum_\sigma \sum_\rho a_\sigma \beta_\rho (U_\mu p_\sigma^\mu, U_\mu p_\rho^\mu)$$

$$= \sum_\sigma \sum_\rho a_\sigma \beta_\rho ([Up_{i_1} \dots Up_{i_\mu}], [Up_{j_1} \dots Up_{j_\mu}])$$

$$= \sum_\sigma \sum_\rho a_\sigma \beta_\rho (Up_{i_1}, Up_{j_1}) \dots (Up_{i_\mu}, Up_{j_\mu})$$

$$= \sum_\sigma \sum_\rho a_\sigma \beta_\rho (p_{i_1}, p_{j_1}) \dots (p_{i_\mu}, p_{j_\mu})$$

$$= \sum_\sigma \sum_\rho a_\sigma \beta_\rho (p_\sigma^\mu, p_\rho^\mu)$$

$$= \left(\sum_\sigma a_\sigma p_\sigma^\mu, \sum_\rho \beta_\rho p_\rho^\mu \right) = (X^\mu, Y^\mu)$$

so U_μ is a unitary operator.

Definition. To each operator L_ϕ^μ with $\mu \leq n$ we associate formally a function ψ_ϕ^μ of $(2n-2\mu)$ variables by considering L_ϕ^μ as a linear mapping of $\mathcal{L}^2(\mathcal{S}^\mu)$ onto itself with $\nu-\mu = m-2\mu$ parameters and letting ψ be the trace of this mapping. Thus, letting $T^\nu = \{T^\mu, \bigwedge^{\nu-\mu}\}$ in (13.8'), we get

$$(13.16)\quad L_\phi^\mu X^\mu = Y^\mu_{\bigwedge^{\nu-\mu}} = \int_{\mathcal{S}^\mu} \phi(S^\mu, T^\mu, \bigwedge^{\nu-\mu}) X^\mu(S^\mu) dS^\mu$$

and

$$(13.17)\quad \psi_\phi^\mu(\bigwedge^{\nu-\mu}) = \mathrm{Tr}\, L_\phi^\mu = \int_{\mathcal{S}^\mu} \phi(S^\mu, S^\mu, \bigwedge^{\nu-\mu}) dS^\mu \quad .$$

Recalling the expression for the integral of a square given on VII-8, we see that in this notation it becomes

$$(13.18)\quad \mathcal{J}(r^2) = \sum_{\mu=0}^{[m/2]} C_{m,\mu} \int_{\mathcal{S}^{m-2\mu}} [\psi_\phi^\mu(\bigwedge^{m-2\mu})]^2 \, d\bigwedge^{m-2\mu}$$

where the $C_{m,\mu}$ are positive integral-valued universal constants. The formal manipulation used to derive this in VII is easily justified in finite dimensional spaces - which is all we require - by direct computation.

We are now ready to prove the

Theorem I. Let $f[\cdot]$ be a functional with degree $m = 2n$ defined on $\mathfrak{H} = \mathcal{L}^2(\mathcal{E})$ by (13.1). Let $L^{\mu}_{\phi,\wedge}$ be defined by (13.16) for almost all $\wedge \in \mathcal{E}^{2n-2\mu}$. We assume that, for each value of μ and for almost all $\wedge$,

$$L^{\mu}_{\phi,\wedge} = A^{+}(\phi,\mu,\wedge) - A^{-}(\phi,\mu,\wedge) \tag{13.19}$$

where A^{+} and A^{-} are positive definite operators on $\mathcal{L}^2(\mathcal{E}^{\mu})$. We further assume that the functions $\psi^{+}_{\phi,\mu}(\wedge)$ and $\psi^{-}_{\phi,\mu}(\wedge)$, defined, for almost all $\wedge \in \mathcal{E}^{2n-2\mu}$, by

$$\psi^{\pm}_{\phi,\mu}(\wedge) = \mathrm{Tr}\, A^{\pm}(\phi,\mu,\wedge) \tag{13.20}$$

are in $\mathcal{L}^2(\mathcal{E}^{2n-2\mu})$.

Then for any basic system of projections (P) the integral $\mathcal{J}_{(P)}(f)$ exists and is independent of the choice of (P). I.e., f is semi-invariant in the sense of V-6.

Proof. We note that in any finite dimensional space the integral of the square of a form fP^k over $\mathfrak{H}$ of degree 2n is given by

$$\mathcal{J}(f^2[P^{\sigma}\cdot]) = \sum_{\mu=0}^{n} C_{n,\mu} \int_{\mathcal{E}^{2n-2\mu}} \left[\int_{\mathcal{E}^{\mu}} \phi^{(\sigma)}(S^{\mu}, S^{\mu}, \wedge^{2n-2\mu}) dS^{\mu} \right]^2 d\wedge^{2n-2\mu} \tag{13.21}$$

and, writing λ for $(2n-2\mu)$ for simplicity,

$$\mathcal{J}(|f[P^{\tau}\cdot] - f[P^{\sigma}\cdot]|^2) = \sum_{\mu=0}^{n} C_{n,\mu} \int_{\mathcal{E}^{\lambda}} \left[\int_{\mathcal{E}^{\mu}} \left\{ \phi^{(\tau)}(S^{\mu}, S^{\mu}, \wedge^{\lambda}) - \phi^{(\sigma)}(S^{\mu}, S^{\mu}, \wedge^{\lambda}) \right\} dS^{\mu} \right]^2 d\wedge^{\lambda} . \tag{13.22}$$

We note that for n = 1 (a quadratic polynomial) these terms reduce to those considered in Chapter VII, pp. 4-7 and, exactly as was noted there (see, again, von Neumann, Math. Found. of Quantum Mech., Ch. II), the assumptions made on ϕ imply, for each μ,

$$(13.23)\qquad \lim_{\sigma,\tau\to\infty}\int_{\mathcal{L}^\lambda}\left[\int_{\mathcal{L}^\mu}\left\{\phi^{(\tau)}(s^\mu,s^\mu,\Lambda^\lambda)-\phi^{(\sigma)}(s^\mu,s^\mu,\Lambda^\lambda)\right\}ds^\mu\right]^2 d\Lambda^\lambda = 0$$

as the integral-ds^μ goes to zero for almost all Λ and the integrand of the integral-$d\Lambda^\lambda$ is bounded for large σ,τ by the integrable function $2[\psi^+_{\phi,\mu}(\Lambda)-\psi^-_{\phi,\mu}(\Lambda)]^2$ and hence goes to zero by the Lebesgue Bounded Convergence Theorem. Therefore

$$(13.24)\qquad \lim_{\sigma,\tau\to\infty}\mathcal{J}(|f[P^\tau\cdot]-f[P^\sigma\cdot]|^2) = 0$$

and f is integrable with respect to (P) for $a = 2$. As the sequence (P) was arbitrary this proves that f is integrable in all coordinate systems.

To show the integral is independent of (P) we note that

$$(13.25)\qquad \mathcal{J}(f[P^k\cdot]) = \mathrm{Tr}\ L^n_\phi(k)$$

which, by the original assumptions with $\mu = n$, must converge to $\mathrm{Tr}\ L^n_\phi$ and so

$$(13.26)\qquad \mathcal{J}_{(P)}(f) = \mathrm{Tr}\ L^n_\phi\ .$$

It is well known (see von Neumann, Math. Found. of Quantum Mech., Ch. II) that if L^n_ϕ is the difference of a pair of positive definite forms, which we have assumed for L^n_ϕ, its trace will be unitarily invariant. Thus $\mathrm{Tr}\ L^n_\phi$ is invariant under all unitary transformations of $\mathcal{L}^2(\mathcal{L}^n)$. Since, by (13.15), we know that a unitary operator on $\mathfrak{H}$ transforms $\mathcal{L}^2(\mathcal{L}^n)$ unitarily, it follows that $\mathcal{J}(f)$ is semi-invariant.

The proof of the invariant integrability of the product of several functions each satisfying the conditions of the previous theorem goes along very similar lines and need only be sketched. By formal manipulation - which, as before, can be verified by direct computation in the finite dimensional case - one sees that the integral of the product and of the square of the difference of two such products are given in terms of integrals in which

variables are paired and then integrated out. For example: in the case of the product of two such functionals, f and g with coefficient functions $\phi(s_1,\ldots,s_\gamma)$ and $\psi(s_1,\ldots,s_\delta)$ where $\gamma+\delta = 2n$, one has

$$(13.27)\quad \mathcal{J}(fg) = \sum_{\mu} C_{\mu}^{n} \int\left[\iint \phi(S^{\mu},S^{\mu},R^{\gamma-2\mu})dS^{\mu}\right]\left[\iint \psi(T^{\bar{\mu}},T^{\mu},R^{\delta-2\bar{\mu}})dT^{\bar{\mu}}\right]dR^{\delta-2\bar{\mu}}$$

where $\gamma-2\mu = \delta-2\bar{\mu}$ and each integral is over the appropriate space. Also

$$(13.28)\quad \mathcal{J}(|(fg)[P^{\tau}\cdot]|^2)$$

$$= \sum_{\mu,\nu,\ldots} C_{\mu,\nu,\ldots} \int\cdots\int\left\{\int\left[\iint \phi^{(\tau)}(S^{\mu},S^{\mu},R^{\nu},T^{\lambda},U^{\rho})dS^{\mu}\int \phi^{(\tau)}(S^{\mu'},S^{\mu'},R^{\nu},T^{\lambda'},U^{\rho'})dS^{\mu'}\right]dR^{\nu}\cdot\right.$$

$$\left.\cdot\int\left[\iint \psi^{(\tau)}(S^{\bar{\mu}},S^{\bar{\mu}},R^{\nu'},T^{\lambda},U^{\rho'})dS^{\bar{\mu}}\int \psi^{(\tau)}(S^{\hat{\mu}},S^{\hat{\mu}},R^{\nu'},T^{\lambda'},U^{\rho})dS^{\hat{\mu}}\right]dR^{\nu'}\right\}dT^{\lambda}dT^{\lambda'}dU^{\rho}dU^{\rho}$$

where
$$2\mu+\nu+\lambda+\rho = 2\mu'+\nu+\lambda'+\rho' = \gamma$$
and
$$2\bar{\mu}+\nu'+\lambda+\rho' = 2\hat{\mu}+\nu'+\lambda'+\rho = \delta \quad .$$

In the expression for $\mathcal{J}(|(fg)[P^{\tau}\cdot]-(fg)[P^{\sigma}\cdot]|^2)$ one has nine sums as in (13.28) with different combinations of τ and σ. We note that if the traces (i.e., the integrations with respect to dS^{μ}, ..., $dS^{\hat{\mu}}$) give square integrable functions - which is guaranteed by the conditions - then the remaining integrations may be expressed as a product of appropriate scalar products by a procedure analogous to (13.14) and all the results will be unitary invariant and continuous in the strong topology. As $\sigma \to \infty$, $P^{\sigma} \to I$ strongly and, by the assumptions, the $\mathcal{L}^2$ differences of approximating traces goes to zero. It follows, as in the first theorem, that

$$(13.29)\quad \mathcal{J}(|(f_1f_2\ldots f_N)[P^{\sigma}\cdot]-(f_1f_2\ldots f_N)[P^{\tau}\cdot]|^2) \to 0$$

and f is integrable with respect to (P). Also, as in the proof of the first theorem it follows that the integral is independent of (P).

It is clear that if f satisfies the requirements of Theorem I it is not only semi-invariant but invariant. To see this we modify the proof of Theorem I as follows: in (13.22) write $\hat{\phi}^{(\sigma)}$ instead of $\phi^{(\sigma)}$ where $\hat{\phi}^{(\sigma)}$ corresponds to Q^{σ} as $\phi^{(\sigma)}$ corresponds to P^{σ}. We note that the integral dS^{μ} in (13.23), with this modification, becomes, for almost all Λ, $\mathrm{Tr}(\{B^{\mu;\sigma,\tau}_{\phi,\Lambda}\}^2)$ where $B^{\mu;\sigma,\tau}_{\phi,\Lambda} = (L^{\mu}_{\phi,\Lambda})P^{\tau} - (L^{\mu}_{\phi,\Lambda})Q^{\sigma}$. Since $L^{\mu}_{\phi,\Lambda}$ is the difference of definite operators, this is of finite double-norm and $B^{\mu;\sigma,\tau}_{\phi,\Lambda} \to 0$ in double-norm as $\sigma,\tau \to \infty$. Hence (see Zaanen, Linear Analysis, p. 244), the integral dS^{μ} goes to zero for almost all Λ and, as with (13.23),

$$(13.30)\quad \lim_{\sigma,\tau\to\infty} \int_{\mathcal{B}^{\lambda}} \left[\int_{\mathcal{B}^{\mu}} \left\{ \phi^{(\tau)}(S^{\mu},S^{\mu},\Lambda^{\lambda}) - \hat{\phi}^{(\sigma)}(S^{\mu},S^{\mu},\Lambda^{\lambda}) \right\} dS^{\mu} \right]^2 d\Lambda^{\lambda} = 0$$

and f is invariant.

Similar considerations apply to products.

Chapter XIV: The Feynman-Kac Formula

As is well known, the formula

$$(14.1) \qquad f(t,x) = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{+\infty} e^{-\frac{(x-y)^2}{2t}} f(y)dy$$

expresses the solution of the initial value problem for the parabolic partial differential equation of "heat flow":

$$(14.2) \qquad \frac{\partial}{\partial t} f(x,t) = \frac{1}{2} \frac{\partial^2}{\partial x^2} f(x,t) \quad , \qquad t \geq 0$$

$$f(x,0) = f(x)$$

To establish this, we differentiate formula (14.1) under the sign of integration

$$(14.3) \qquad \frac{\partial}{\partial t} f(x,t) = -\frac{1}{2} t^{-1} f(x,t)$$

$$+ \frac{1}{2} t^{-2} \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{+\infty} (x-y)^2 e^{-\frac{(x-y)^2}{2t}} f(y)dy$$

$$\frac{\partial}{\partial x} f(x,t) = -\frac{1}{\sqrt{2\pi t}} t^{-1} \int_{-\infty}^{+\infty} (x-y) e^{-\frac{(x-y)^2}{2t}} f(y)dy$$

$$\frac{\partial^2}{\partial x^2} f(x,t) = -t^{-1} f(x,t) + t^{-2} \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{+\infty} (x-y)^2 e^{-\frac{(x-y)^2}{2t}} f(y)dy \ ,$$

proving the validity of the partial differential equation (14.2) for each function f such that

$$\int_{-\infty}^{+\infty} e^{-\varepsilon x^2} |f(x)|dx < \infty$$

for each $\varepsilon > 0$. To establish the validity of the initial condition (14.2) we must suppose, in addition, that $f(x)$ is continuous; in this case, the argument required is a standard 'singular integral' argument the details of which are left to the reader as an exercise.

Formula (14.1) may readily be written as a Gaussian integral. Indeed, we have

$$(14.4)\quad \int_{\mathcal{H}} f\left(\int_0^t \phi(s)ds + x\right)\mu_G(d\phi) = \int_{\mathcal{H}} f\left((\psi_t,\phi)+x\right)\mu_G(d\phi)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f\left(x+t^{1/2}s\right) e^{-1/2s^2}ds$$

$$= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{+\infty} f(x+y)e^{-y^2/2t}\,dy$$

$$= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{+\infty} f(u)e^{-\frac{(x-u)^2}{2t}}\,du$$

where the Hilbert space $\mathcal{H}$ is taken to be the space $L_2[0,\infty)$ of square-integrable functions, and where ψ_t denotes the characteristic function of the interval $[0,t]$. The computations are valid since the integrand of the Gaussian integral on the right-hand side of this formula is a cylinder functional whose integral may be evaluated by formula (3.1) of Chapter V.

A corresponding integral formula is known for the solution of the more general parabolic initial value problem.

$$(14.5)\quad \frac{\partial}{\partial t}F(x,t) = \frac{1}{2}\frac{\partial^2}{\partial x^2}F(x,t) + V(x,t)F(x,t)\ ,\ t \geq 0\ ;$$

$$F(x,0) = F(x)\ ,$$

V being a given coefficient function. This formula is

$$(14.6)\qquad F(x,t) = \int_{\mathcal{Q}} F\left(\int_0^t \phi(s)\,ds + x\right) \exp\left[\int_0^t V\left(s, \int_s^t \phi(u)\,du + x\right) ds\right] \mu_G(d\phi) .$$

The main point of the present chapter is to establish this formula. We shall divide the proof into the proof of two lemmas, which will be stated in terms of the following definition.

(14.7) Definition: Let $V(t,x)$ be a given "coefficient function", defined for all x and all $t \geq 0$, and continuous and uniformly bounded in t and x in its range of definition. Let $F(x)$ be an integrable function of the real variable x; then, for $0 \leq s < t$, we define the mapping $T_{s,t}$ by the formula

$$(14.8)\qquad [T_{s,t}F](x) = \int_{\mathcal{Q}} F\left(\int_s^t \phi(u)\,du + x\right) \exp\left[\int_s^t V\left(u, \int_u^t \phi(v)\,dv + x\right) du\right] \mu_G(d\phi).$$

$$(14.9)\quad \text{Lemma:}\quad \frac{d}{dt}\,[T_{s,t}F](x)\Big|_{t=s} = \frac{1}{2}\,\frac{\partial^2}{\partial x^2} F(x) + V(x,s)F(x)$$

if the function F is integrable and has two continuous derivatives, and if $s > 0$.

Proof of Lemma 14.9: It follows immediately from definition (14.7) that for each integrable function F the quantity $[T_{t,t_1}T_{s,t}F](x)$ may be written as an iterated Gaussian integral as follows:

$$(14.11)\quad [T_{t,t_1}T_{s,t}F](x) = \int_{\mathcal{Q}}\int_{\mathcal{Q}} F\left(\int_s^t \phi(u)\,du + \int_t^{t_1} \psi(u)\,du + x\right) \cdot$$
$$\cdot \exp\left[\int_s^t V\left(u, \int_u^t \phi(v)\,dv + \int_t^{t_1} x(v)\,dv + x\right) du\right] \cdot$$
$$\cdot \exp\left[\int_t^{t_1} V\left(u, \int_u^{t_1} \psi(v)\,dv + x\right) du\right] \mu_G(d\phi)\mu_G(d\psi).$$

Next, it must be observed that a striking form of Fubini's theorem holds for Gaussian integration over Hilbert space. Suppose that $\mathfrak{H}_1$ and $\mathfrak{H}_2$ are two Hilbert spaces. Then their direct sum $\mathfrak{H}_1 \oplus \mathfrak{H}_2$ is also a Hilbert space, and hence, if $\mathfrak{H}_3$ is any other Hilbert space (of the same dimension) there exists an isometric mapping U of $\mathfrak{H}_1 \oplus \mathfrak{H}_2$ onto $\mathfrak{H}_3$. If $[\phi,\psi]$ is a pair of elements such that $\phi \in \mathfrak{H}_1$, $\psi \in \mathfrak{H}_2$, then $\eta = U[\phi,\psi]$ will be an element of $\mathfrak{H}_3$. Let $G(\phi,\psi)$ be a function of the two variables ϕ and ψ; suppose first that $G(\phi,\psi)$ is a cylinder function on $\mathfrak{H}_1 \oplus \mathfrak{H}_2$, so that there exist orthonormal elements $\{\phi_1,\dots,\phi_n\}$ of $\mathfrak{H}_1$, orthonormal $\{\psi_1,\dots,\psi_m\}$ of $\mathfrak{H}_2$, and a function g of (n+m) real variables, such that

$$(14.12)\qquad G(\phi,\psi) = g\Big((\phi,\phi_1),\dots,(\phi,\phi_n),(\psi,\psi_1),\dots,(\psi,\psi_m)\Big).$$

Then, by formula (3.1) of Chapter V,

$$(14.13)\qquad \int_{\mathfrak{H}_1}\int_{\mathfrak{H}_2} G(\phi,\psi)\mu_G(d\phi)\mu_G(d\psi)$$

$$= \sqrt{2\pi}^{-(n+m)} \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} g(x_1,\dots,x_{n+m})e^{-\frac{1}{2}(x_1^2+\dots+x_{n+m}^2)}dx_1\cdots dx_{n+m}$$

On the other hand, let $\mathscr{G}(\eta) = G(U^{-1}\eta)$. Then $\{U[\phi_1,0],\dots U[\phi_n,0],$ $U[0,\psi_1]\dots,U[0,\psi_m]\}$ form an orthonormal set in $\mathfrak{H}_3$, which we may denote as $\{\eta_1,\dots,\eta_{n+m}\}$, and clearly

$$(14.14)\qquad \mathscr{G}(\eta) = g\Big((\eta,\eta_1),\dots,(\eta,\eta_{n+m})\Big).$$

By formula (3.1) of Chapter V, $\int_{\mathfrak{H}_3}\mathscr{G}(\eta)\mu_G(d\eta)$ is expressed by the (n+m)-fold Lebesgue integral on the right of (14.13).

Thus we find that

$$\int_{\mathfrak{H}_1}\int_{\mathfrak{H}_2} G(\phi,\psi)\mu_G(d\psi) = \int_{\mathfrak{H}_3} \mathcal{G}(\eta)\mu_G(d\eta),$$

i.e., that

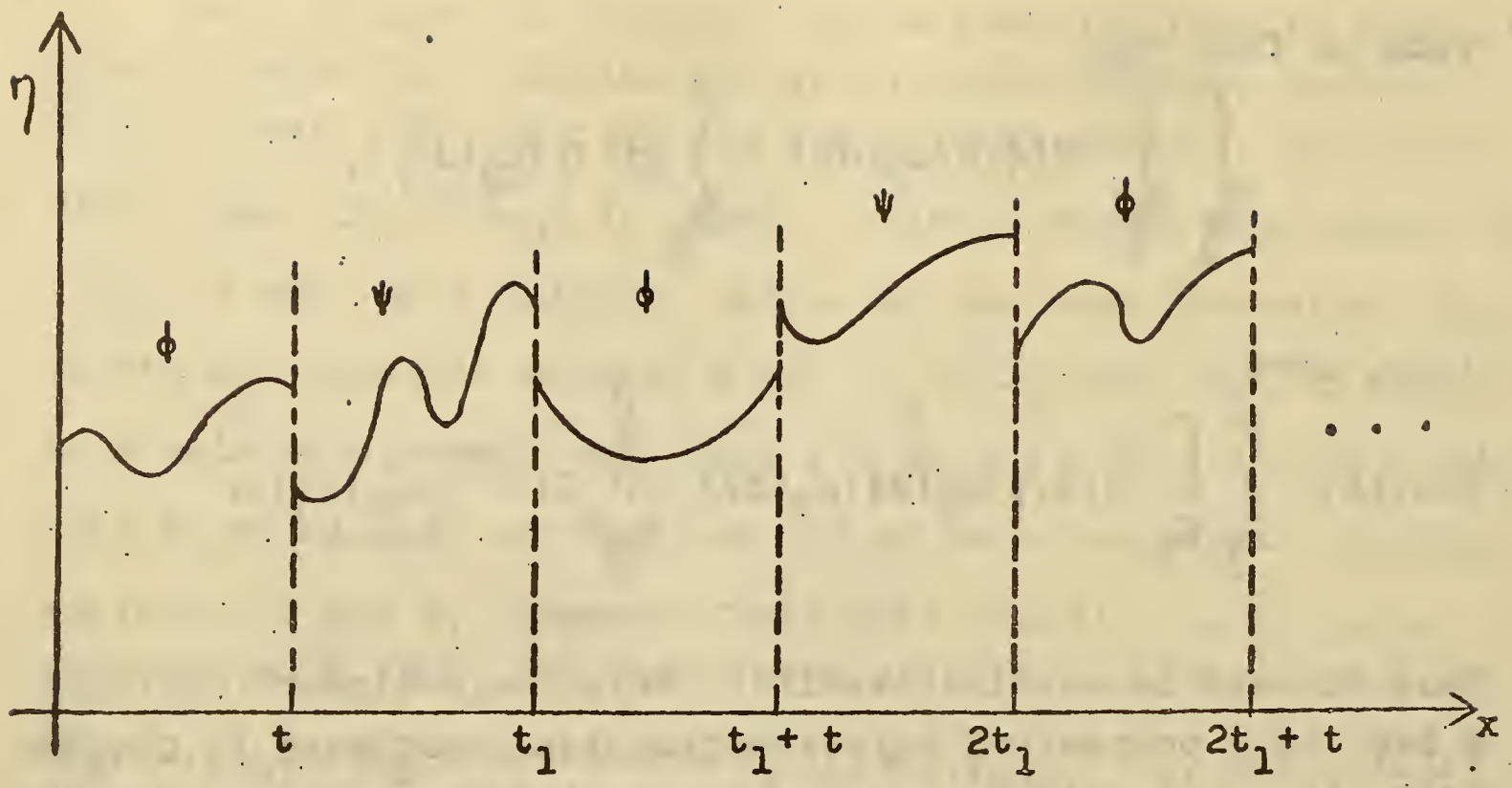
$$(14.15)\qquad \int_{\mathfrak{H}_1}\int_{\mathfrak{H}_2} G(\phi,\psi)\mu_G(d\phi)\mu_G(d\psi) = \int_{\mathfrak{H}_3} G(U^{-1}\eta)\mu_G(d\eta)\ .$$

This formula is at first verified only for cylinder functions G but, by a process of approximation like that used in Chapter VIII, it may readily be shown to be valid for all functions G integrable in the sense of that chapter. The details of this argument are left to the reader as an exercise. We wish to consider the application of formula (14.15) to the proof of Lemma 14.9 . Suppose then that the Hilbert spaces $\mathfrak{H}_1$ and $\mathfrak{H}_2$ are both the space $L_2[0,\infty)$. Then the space $\mathfrak{H}_3$ may also be taken to be $L_2[0,\infty)$, and the isometry U may be constructed as follows in terms of the quantities t and t_1 of formula (14.11).

Let $\phi \in L_2[0,\infty)$ and $\psi \in L_2[0,\infty)$. Define $\eta = U[\phi,\psi]$ by

$$\eta(x) = \begin{cases} \phi(x) & 0 \le x \le t \\ \psi(x-t) & t < x \le t_1 \\ \phi(x-t_1-t) & t_1 < x \le t_1+t \\ \psi(x-2t) & t_1+t < x \le 2t_1 \\ \phi(x-2t_1-2t) & 2t_1 < x \le 2t_1+t \\ \cdots & \end{cases}$$

A diagram may perhaps make this c nstruction more clear.



Then clearly $$\int_s^t \phi(u)du + \int_t^{t_1} \psi(u)du = \int_s^{t_1} \eta(u)du,$$

$$\int_u^t \phi(v)dv = \int_t^{t_1} \psi(v)dv = \int_u^{t_1} \eta(v)dv \qquad u \in [s,t],$$

and $$\int_u^t \psi(v)dv = \int_u^t \eta(v)dv \qquad u \in [t,t_1].$$

Thus, using formula (14.11) and formula (14.15), we find that

(14.16) $$[T_{t,t_1} T_{s,t} F](x)$$

$$= \int_{\mathcal{H}} F\left(\int_s^{t_1} \eta(u)du + x\right) \exp\left[\int_s^t V\left(u, \int_u^{t_1} \eta(v)dv + x\right) du\right] \cdot$$

$$\cdot \exp\left[\int_t^{t_1} V\left(u, \int_u^{t_1} \eta(v)dv + x\right) du\right] \mu_G(d\eta)$$

$$= \int_{\mathcal{H}} F\left(\int_s^{t_1} \eta(u)du + x\right) \exp\left[\int_s^{t_1} V\left(u, \int_u^{t_1} \eta(v)dv + x\right) du\right] \mu_G(d\eta)$$

so $$[T_{t,t_1}T_{s,t}F](x) \equiv [T_{s,t_1}F](x)$$

proving Lemma 14.9 . Q. E. D.

Proof of Lemma 14.10: Let s be given, and, for each function F satisfying the hypotheses of Lemma 10.10, let $W_{t,t_1,t_2}F$ be defined by

$$(14.17)\quad [W_{t,t_1,t_2}F](x) = \int_{\mathcal{S}'} F\left(\int_s^t \phi(u)du+x\right) \exp\left[\int_s^{t_1} V\left(u,\int_u^{t_2} \phi(v)dv+x\right) du\right]\mu_G(d\phi).$$

Then

$$(14.18)\quad [W_{t,s,t_2}F](x) = \int_{\mathcal{S}'} F\left(\int_s^t \phi(u)du+x\right) \mu_G(d\phi),$$

so that, by the first paragraph of the present section,

$$(14.19)\quad \frac{\partial}{\partial t}[W_{t,s,t_2}F](x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}[W_{t,s,t_2}F](x) .$$

Thus, putting t = s, it follows, since $[W_{s,s,t_2}F](x) = F(x)$ by (14.18) and by the first paragraph of the present section, that

$$(14.20)\quad \frac{\partial}{\partial t}[W_{t,s,t_2}F](x)\Big|_{t=s} = \frac{1}{2}\frac{\partial^2}{\partial x^2}F(x)$$

Moreover,

$$(14.21)\quad \frac{\partial}{\partial t_1}[W_{t,t_1,t_2}F](x) = \int_{\mathcal{S}'} F\left(\int_s^t \phi(u)du+x\right)\cdot$$

$$\cdot \exp\left[\int_s^{t_1} V\left(u,\int_u^{t_2} \phi(v)dv+x\right)du\right] V(t_1,x)\mu_G(d\phi)$$

$$= V(t_1,x)[W_{t,t_1,t_2}F](x),$$

while evidently

$$\frac{\partial}{\partial t_2}\,[W_{t,s,t_2}F](x) = 0 \,. \tag{14.22}$$

Thus

$$\frac{\partial}{\partial t}\,[W_{t,t,t}F](x)\Big|_{t=s} = \frac{\partial}{\partial t}\,[W_{t,s,s}F](x)\Big|_{t=s} = \frac{\partial}{\partial t}[W_{s,t,s}F](x)\Big|_{t=s}$$

$$= \frac{\partial}{\partial t}[W_{s,t,s}F](x)\Big|_{t=s} + \frac{\partial}{\partial t}[W_{s,s,t}F](x)\Big|_{t=s}$$

$$= \frac{1}{2}\,\frac{\partial^2}{\partial x^2}\,F(x)\;V(s,x)F(x)\;,$$

since $\frac{\partial}{\partial t}\,[W_{t,s_1,t_2}F](x)$ and $\frac{\partial}{\partial t_1}\,[W_{t,t_1,t_2}F](x)$ are readily seen from (14.19) and (14.21) to be continuous in t, t_1, and t_2 so that the chain rule for differentiation is valid. Since $[W_{t,t,t}F](x) = [T_{s,t}F](x)$, this proves Lemma 14.10 .

Now, combining Lemmas 14.9 and 14.10 we find that

$$\frac{\partial}{\partial t}\,[T_{0,t}F](x) = \frac{\partial}{\partial s}\,[T_{t,s}T_{0,t}F](x)\Big|_{s=t} \tag{14.23}$$

$$= \left\{\frac{1}{2}\,\frac{\partial^2}{\partial x^2} + V(t,x)\right\}\,[T_{0,t}F](x)\;.$$

Thus $[T_{0,t}F](x) = F(x,t)$ defines a solution of the parabolic partial differential equation (14.5). Since

$$[T_{0,0}F](x) = \int_{\mathcal{C}} F\left(\int_0^t \phi(s)ds + x\right) \exp\left[\int_0^t V\left(s, \int_s^t \phi(u)du + x\right) ds\right] \mu_G(d\phi)\Big|_{t=0}$$

$$= \int_{\mathcal{C}} F\left(\int_0^t \phi(s)ds + x\right)\mu_G(d\phi)\Big|_{t=0} = F(x)\;,$$

we have also $F(x,0) = F(x)$. Thus the fact that the integral formula (14.6) defined a solution of the parabolic initial value problem (14.5) is completely verified.

A corresponding formula may readily be written for the solution of the more general initial value problem for the parabolic partial differential equation

$$(14.24)\qquad \frac{\partial}{\partial t} F(x_1,\dots,x_n;t) = \frac{1}{2}\frac{\partial^2}{\partial x_1^2} + \dots + \frac{\partial^2}{\partial x_n^2} F(x_1,\dots,x_n;t)$$

$$t \geq 0;$$

$$+ V(x_1,\dots,x_n;t)F(x_1,\dots,x_n;t)$$

$$F(x_1,\dots,x_n;0) = F(x_1,\dots,x_n)\ ,$$

V being a given coefficient function. To write this formula we let $\mathfrak{H}$ denote the space of all functions $\phi(s)$ of the real variable s, defined for $s \geq 0$, and having values in the space E^n of n-dimensional vectors, such that

$$\int_0^\infty ||\phi(s)||^2 ds = ||\ \phi\ ||^2 < \infty$$

This space is a Hilbert space so that a Gaussian integral is defined for functions on $\mathfrak{H}$. If we write $F(x)=F(x_1,\dots,x_n)$, where x is the vector in E^n with components $x_1,\dots x_n$, then the desired formula for the function F(x,t) of (14.24) can be written as

$$(14.25)\qquad F(x;t)=\int_{\mathfrak{H}} F\left(\int_0^t \phi(s)ds+x\right) \exp\left[\int_0^t V\left(s,\ \int_s^t \phi(u)du+x\right)ds\right]\mu_G(d\phi),$$

a formula exactly similar in form to (14.6). The proof in this new case is quite like the proof that formula (14.6) defines a solution to the parabolic initial value problem (14.5). In

the special case $V \equiv 0$, in which (14.24) refers to the n-dinensional heat equation, it is readily seen that the integral (14.25) is

$$(14.26)\quad F(x_1,\dots,x_n;t)=F(x;t)=\int_{\mathfrak{F}} F\left(\int_0^t \phi(s)ds+x\right)\mu_G(d\phi)$$

$$=\sqrt{2\pi}^{-n}\int^{n} F(y_1,\dots,y_n)\exp\left[-\frac{1}{2t}\left\{(x_1-y_1)^2+\dots+(x_n-y_n)^2\right\}\right]\cdot dy_1\dots dy_n,$$

i.e., is the ordinary integral for the solution of the n-dimensional heat equation. The proof of the general formula (14.25) uses the results known in this special case. To the reader is left all elaboration of the details of this proof.

Formulae (14.6) and (14.25) provide a powerful means for calculating certain non-elementary Gaussian integrals in Hilbert space. Suppose, for instance, that V is a given measurable function of the real variable s, that e is a Borel set of real numbers, and that we wish to calculate the Gaussian measure $\mu_V(\tilde{e})$ of the subset $\tilde{e}$ of Hilbert space defined by

$$\tilde{e} = \left\{\phi\in\mathfrak{F}: \int_0^1 V\left(\int_0^t \phi(u)du\right)dt\in e\right\}.$$

This problem may be solved as follows. From the definition of an integral as a limit of sums we have evidently

$$(14.27)\quad \int_{-\infty}^{+\infty} e^{i\lambda t}\,\mu_V(dt) = \int_{\mathfrak{F}} \exp\left[i\lambda\int_0^1 V\left(\int_0^t \phi(u)du\right)dt\right]\mu_G(d\phi).$$

The integral on the right of (14.27) may be expressed, using (14.6), as $F_\lambda(0;1)$, where $F_\lambda(x;t)$ is the solution of the initial-value problem

$$(14.28)\quad \frac{\partial}{\partial t}F_\lambda(x,t) = \frac{1}{2}\frac{\partial^2}{\partial x^2}F_\lambda(x,t) + i\lambda V(x)F_\lambda(x,t) \qquad t\le 0;$$

$$F_\lambda(x,0)\equiv 1$$

If the solution of this initial-value problem can be given explicitly, then inversion of the Fourier transform on the left of (14.27) will give an explicit expression for μ_V.

The Gaussian integral with respect to the"general Gaussian measure" defined by formula (9.1) of Chapter V may be used to express the solution of the initial value problem for the parabolic P.D.E.

$$(14.29)\quad \frac{\partial}{\partial t} F(x,t) = \frac{1}{4\sigma} \frac{\partial^2}{\partial x^2} F(x,t) + V(x,t)F(x,t)$$

$$t \leq 0;$$

$$F(x,0) = F(x)$$

as

$$(10.30)\quad F(x,t)=\int_{\mathcal{S}_w} F\left(\int_0^t \phi(s)ds+x\right)\exp\left\{\int_0^t V\left(s,\int_s^t \phi(u)du+x\right)ds\right\}\mu_G^{\sigma}(d\phi),$$

where μ_G^{σ} denotes the Gaussian measure defined by the right side of formula (9.1) of Chapter V. If, as in Chapter V, we allow σ to be complex, we can hope to express the solution of an even more general class of partial differential equations as integrals over Hilbert space. In particular, if we go "to the limiting case" relative to the considerations of §§ 9 of Chapter V, and take $\sigma = -\frac{1}{4}\,i$, we obtain (at least formally) an integral over Hilbert space which expresses the solution of the "Schrödinger equation"

$$(10.31)\qquad \frac{1}{i}\,\frac{\partial}{\partial t} F(x,t) = \frac{\partial^2}{\partial x^2} F(x,t) + V(x,t)F(x,t)$$

$$F(x,0)= F(x)$$

This was the form in which the class of formulae treated in the present chapter was first proposed by the physicist Feynman, in connection with certain researches in quantum theory. Because of the still unresolved analytical difficulties in treating such complex Gaussian measures over Hilbert space, we shall not pursue this matter to any greater extent.

Chapter XV

The Wiener-Siegel Causal Theory of Quantum Mechanics

The ordinary formalism of quantum theory may be described in hurried outline as follows. Each observable property of a given physical system is associated with an Hermitian operator A on a (complex) Hilbert space $\mathfrak{H}$. The evolution of the physical system in time brings about a systematic and continuous change of these observables and a corresponding change of the associated Hermitian operators: this latter change is determined by the Heisenberg-Schrödinger equation

$$A(t) = e^{i(t-t_o)H} A(t_o) e^{-i(t-t_o)H} ,$$

which may be written in differential form as

$$-i \frac{d}{dt} A(t) = HA(t) - A(t)H ,$$

here H is the so-called "energy operator" for the system. In the simplest cases, the physical system will be in a certain "pure statistical state" determined by a vector $x \in \mathfrak{H}$ of norm 1. In such a state, each observable has associated with it a certain well-determined probability distribution, which gives the probability that a measurement of the observable corresponding to the operator A will lie in any given subset of the real axis. The fundamental statistical law of quantum theory states that the expected value, EXP(A) or average for the measurement of the observable corresponding to A, when the given system is in a statistical state determined by the vector x is given by

$$\text{EXP}(A) = (Ax,x) . \tag{15.1}$$

Now, the problem of a causal interpretation of quantum theory, in general, is to recast formula (15.1) into the form of an ordinary statistical assertion. More precisely, one wants to find

(15.2)

(a) a space Ω of elements ω. These elements ω are to be thought of heuristically as "hidden variables", knowledge of which would suffice to determine the "true values" of the observable A, not probabilistically, but definitely. Thus, with each observable A one wants to associate

(b) a function $f_A(\omega)$ defined on the space Ω. The next step would be to interpret (15.1) as a true statistical formula; that is, to find

(c) a probability-measure μ_x on the space Ω, associated with each state vector $x \in \mathfrak{H}$ in such a way that

$$\int_{\Omega} f_A(\omega)\mu_x(d\omega) = \text{EXP } (A) = (Ax,x) \ .$$

Von Neumann showed that, in case the correspondence $A \rightarrow f_A$ is assumed to have a number of reasonable algebraic properties (such as linearity, etc.), the problem of hidden variables as defined by (15.2) has no solution. N. Wiener and A. Siegel, in two recent papers on this question, have made an observation in this connection which it is the aim of the present chapter to explain. (See the second of these two papers, Nuovo Cimento Ser. X, v. II, p. 982, (1955) for a reference to the first paper, and for what follows.)

It should first of all be noted that the Hilbert space $\mathfrak{H}$ involved here is complex, rather than real as has been the case till now. Thus, if $\{\phi_n\}$ is an orthonormal basis for $\mathfrak{H}$, each vector $x \in \mathfrak{H}$ can be written uniquely as $x = \sum_{i=1}^{\infty} a_i \phi_i$, where $\{a_i\}$ is a sequence of <u>complex</u> numbers such that $\sum_{i=1}^{\infty} |a_i|^2 < \infty$, rather than a sequence of real numbers such that $\sum_{i=1}^{\infty} a_i^2 < \infty$ as has been the case till now. The simplest way to define a Gaussian integral in the complex Hilbert space $\mathfrak{H}$ is as follows. Let $\mathfrak{H}_1$ be a real Hilbert space, with an orthonormal basis $\{\psi_n\}$.

Write a complex number a in terms of its real and imaginary parts as $a = a + ib$. Then the mapping U defined by

$$(15.3)\quad U(a_1\phi_1 + a_2\phi_2 + \ldots) = (a_1\psi_1 + b_1\psi_2 + a_2\psi_3 + b_2\psi_4 + \ldots)$$

determines an isometry of $\mathfrak{h}$ onto $\mathfrak{h}_1$. If f is a function defined on $\mathfrak{h}$, we call f Gauss-integrable if the function $f(U^{-1}(\cdot))$, which is a function defined on $\mathfrak{h}_1$, is Gauss-integrable, and write

$$(15.4)\quad \int_{\mathfrak{h}} f(x)\mu_G(dx) = \int_{\mathfrak{h}_1} f(U^{-1}(y))\mu_G(dy)\ ,$$

defining the integral on the left by that on the right. If V is a unitary transformation of $\mathfrak{h}$ onto itself, it is clear that $V_1 = UVU^{-1}$ is a unitary transformation of $\mathfrak{h}_1$ onto itself. Thus

$$(15.5)\quad \int_{\mathfrak{h}} f(Vx)\mu_G(dx) = \int_{\mathfrak{h}_1} f(VU^{-1}y)\mu_G(dy)$$

$$= \int_{\mathfrak{h}_1} f(U^{-1}V_1y)\mu_G(dy)$$

$$= \int_{\mathfrak{h}_1} f(U^{-1}y)\mu_G(dy)$$

$$= \int_{\mathfrak{h}} f(x)\mu_G(dx)\ ,$$

proving that the Gaussian integral in complex Hilbert space is also unitary invariant. If f is a cylinder function on $\mathfrak{h}$, which we write as $f(x) = \phi((x,v_1),\ldots,(x,v_n))$, $\{v_1,\ldots,v_n\}$ being an orthonormal set of vectors, and $\phi(z_1,\ldots,z_n)$ a function of n complex variables, it follows readily from definition (15.4) that

$$(15.6)\quad \int_{\mathfrak{h}} f(x)\mu_G(dx) = (2\pi)^{-n}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\phi(a_1+ib_1,\ldots,a_n+ib_n)$$

$$\cdot \exp\left[-\tfrac{1}{2}(a_1^2 + b_1^2 + \ldots + a_n^2 + b_n^2)\right]da_1\ldots db_n$$

$$= (2\pi)^{-n}\int_0^{+\infty}\cdots\int_0^{+\infty}\int_0^{2\pi}\cdots\int_0^{2\pi}\phi(r_1e^{i\theta_1},\ldots,r_ne^{i\theta_n})$$

$$\cdot \exp\left[-\tfrac{1}{2}(r_1^2+\ldots+r_n^2)\right]d\theta_1\ldots d\theta_n r_1dr_1\ldots r_ndr_n\ .$$

This last form, which gives the integral in "polar coordinates", will be useful below.

We are now ready to state the Wiener-Siegel formula in its connection with the problem defined by (15.2). Let the physical system be in the statistical state determined by the vector x. As the space Ω of hidden variables we simply take the Hilbert space $\mathfrak{H}$ itself. As the measure of (15.2) (c) we take the Gaussian measure μ_G in Hilbert space. Thus Wiener and Siegel's measure μ_G is independent of the state vector x. The function f_A of (15.2) (b) is defined for a Hermitian operator A as follows, at any rate in the special (but not terribly special) case in which A has a discrete spectrum with a complete set $\{\xi_n\}$ of eigenvectors and a corresponding set $\{\lambda_n\}$ of real eigenvalues:

(15.7) If there is an n such that, for every $m \neq n$,

$$|(y,\xi_n)/(x,\xi_n)|^2 < |(y,\xi_m)/(x,\xi_m)|^2$$

then put

$$f_A(y) = \lambda_n .$$

Otherwise, put

$$f_A(y) = 0 .$$

(The set of y involved in the second case is a set of Gaussian measure zero, which is why it may be treated so cavalierly.)

Since the function f_A defined by (15.7) depends also on the state-vector x, it is better to write it as $f_A^{(x)}(y)$. At any rate, the fundamental statistical formula (15.2) (c) which must now be proved is

$$(15.8) \qquad \int_{\mathfrak{H}} f_A^{(x)}(y)\mu_G(dy) = (Ax,x) .$$

To prove this, suppose that we make use of the unitary invaraince of the Gaussian integral to write Hilbert space in a representation $\{a_1,a_2,\ldots\}$ in which A is diagonal, i.e., in which $A\{a_1,a_2,\ldots\} = \{\lambda_1 a_1, \lambda_2 a_2, \ldots\}$. Then, what must be shown is

$$(15.9)\quad \sum_{j=1}^{\infty} \lambda_j x_j^2 = (Ax,x) = \int_{\Lambda} f_A^{(x)}(y)\mu_G(dy)$$

$$= \sum_{j=1}^{\infty} \lambda_j (2\pi)^{-\infty} \left(\int_0^{2\pi} d\vartheta_j\right)^{\infty} \int_0^{\infty} \exp[-t_j]dt_j \left\{\prod_{r\neq j} \int_{X_{j,r}}^{\infty} \exp[-t_r]dt_r\right\}$$

where $X_{j,r} = t_j x_r^2/x_j^2$. Here $x_j = |(x,\xi_j)|$ by definition; we have made the substitution $t_j = \frac{1}{2}r_j^2$ in the polar coordinate form (15.6) of the integral (15.8), and have also made use of the definition (15.7) of the function $f_A^{(x)}$. Thus, what must be shown is that

$$(15.10)\quad x_j^2 = \int_0^{\infty} \exp[-t_j]dt_j \cdot \left\{\prod_{r\neq j} \int_{X_{j,r}}^{\infty} \exp[-t_r]dt_r\right\}.$$

But we have

$$(15.11)\quad \int_0^{\infty} \exp[-t_j]dt_j \left\{\prod_{r\neq j} \int_{X_{j,r}}^{\infty} \exp[-t_r]dt_r\right\}$$

$$= \int_0^{\infty} \exp[-t_j] \prod_{r\neq j} \exp\{-t_j x_r^2/x_j^2\}$$

$$= \int_0^{\infty} \exp\left[-t_j\left(1 + \sum_{r\neq j} x_r^2/x_j^2\right)\right]dt_j$$

$$= \int_0^{\infty} \exp\{-t_j/x_j^2\}dt_j = x_j^2 .$$

Here we have used the fact that $\|x\| = 1$ in the form $\sum_{r=1}^{\infty} x_r^2 = 1$. Thus (15.10) and (15.9) are established.

So much for the mathematical side of the present discussion. It should be noted on the physical side that in the statement (15.2) of the problem of hidden variables it was <u>essential</u> that the function f_A defining the true value of the observable corresponding to the hermitian operator A at the various points in phase space Ω was <u>independent of the statistical state</u> x and that this statistical state x determined <u>only the measure</u> μ_x on the space Λ. These requirements spring directly from the very notion of a "statistical interpretation", in the classical sense,

of a physical process, since this notion demands that the observed average values of observables should be defined as a 'harmless' average of completely non-statistical mechanical quantities, well defined as functions on a phase space, without any reference to statistical considerations. To allow the very functions on phase space which define the values of observable quantities to depend on a 'statistical state' is thus to offer a theory that is very far indeed from a classical statistical theory, and hence from anything that can justifiably be called a 'causal' or 'hidden variables' theory of quantum mechanics. For this reason it may be considered that, in a strict sense, the results of Wiener and Siegel do not yield a truly causal theory of quantum mechanics.

Bibliographic Appendix

The specific approach to integration in Hilbert space developed in these notes was initiated by K.O. Friedrichs in [1], p. 52-63. Most of the earlier work on related problems concerned the Wiener integral discussed in Chapter IX above. The fundamental early papers were Wiener [1] and [2], and Paley, Wiener, and Zygmund [1]. Most of the subsequent progress is due to R.H. Cameron, W.T. Martin, and their students. Cameron and Marton [3], [5], [7], and Cameron and Fagan [1] study the effect of various sorts of mappings in the space of continuous functions on the Wiener measure, establishing generalizations of the Jacobian determinant law governing changes of variable in finite dimensional multiple integrals. Cameron and Martin [6] evaluate various Wiener integrals by the use of a general method involving 'changes of variable' in the space of continuous functions, and making use of certain Sturm-Liouville differential equations. Cameron and Martin [2], Cameron and Hatfield [1], [2] establish a complete set of "Fourier-Hermite functionals" in the space of functions square-integrable over the Wiener measure space, and study the associated series-expansion theory, Cameron [4] and Cameron and Martin [9] develop a related theory of Fourier-Wiener integrals. Cameron and Martin [1], [4], Cameron and Shapiro [1], Cameron, Lindgren, and Martin [1] show how the solutions of certain nonlinear integral equations can be expressed 'explicitly' in terms of Wiener integrals. Cameron [2] and Owchar [1] study formulae for Wiener integrals related to the formula

$$\frac{d}{dt}\int_0^t f(x)dx = f(t)$$

for functions of a real variable. Cameron [1] gives a

'Simpson's Rule' for the numerical evaluation of Wiener integrals. See also Kac [2], Maruмaya [1], Cameron and Graves [1], Cameron [4], and Wiener [3], Paley and Wiener [1], and Chapter IX above.

The development of the Feynman-Kac theory presented in Chapter X above begins with the thesis of the physicist Feynman [1], and continues in the studies of Kac [1] and [3], Rosenblatt [1], Tingley [1], Fortet [1], Cameron [3], Montroll [1], Steinberg [1]. The authors cited all investigate various aspects of the relationship between a partial differential equation of parabolic type and the Wiener-integral formula for its solution. Cameron [3] in particular gives very detailed results under very mild analytical restrictions. It should be remarked that the relationship cited is a particular case of a more general relationship, known from the probability theory of Markoff processes, between a Markoff process in its 'transition probability' and its 'sample function' representations. For applications of the Feynman-Kac formula to the theory of probability, see Kac [1] and Erdos and Kac [1]. Gelfand and Yaglom [1] give an excellent review of the theory of the Wiener integral with special emphasis on the Feynman-Kac formula.

For various other, quite different, approaches to the problem of integration in Hilbert space see K. Löwner [1], E.R. Lorch [1], [2], [3], Friedrichs [1], p. 121-132, P. Levy [1], p.209-355.

BIBLIOGRAPHY:

Cameron, R.H.

1. A "Simpson's Rule" for the numerical evaluation of Wiener's integrals in function space. Duke Math. J. 18, 111-130 (1951).
2. The first variation of an indefinite Wiener integral. Proc. Amer. Math. Soc. 2, 914-924 (1951).
3. The generalised heat flow equation and a corresponding Poisson formula. Ann. of Math. (2) 59, 434-462 (1954).
4. Some examples of Fourier-Wiener transforms of analytic functionals. Duke Math. J. 12, 485-488 (1945).
5. The translation pathology of Wiener space. Duke Math. J. 21, 623-627 (1954).

Cameron, R.H. and Fagan, R.E.

1. Non-linear transformations of Volterra type in Wiener space. Trans. Amer. Math. Soc. 75, 552-575 (1953).

Cameron, R.H. and Graves, R.E.

1. Additive functionals on a space of continuous functions. I. Trans. Amer. Math. Soc. 70, 160-176 (1951).

Cameron, R.H. and Hatfield, C.

1. Summability of certain orthogonal developments of non-linear functionals. Bull. Amer. Math. Soc. 55, 130-145 (1949).
2. Summability of certain series of unbounded non-linear functionals. Proc. Amer. Math. Soc. 4, 375-387 (1953).

Cameron, R.H., Lindgren, B.W., and Martin, W.T.

1. Linearization of certain non-linear functional equations. Proc. Amer. Math. Soc. 3, 138-143 (1952).

Cameron, R.H. and Martin, W.T.

1. An expression for the solution of a class of non-linear integral equations. Amer. J. Math. 66, 281-298 (1944).
2. The orthogonal development of non-linear functionals in series of Fourier-Hermite functionals. Ann. of Math. (2) 48, 385-392 (1947).
3. The transformation of Wiener integrals by non-linear transformations. Trans. Amer. Math. Soc. 66, 253-283 (1949).
4. Non-linear integral equations. Ann. of Math. (2) 51, 629-642 (1950).
5. Transformations of Wiener integrals under a general class of linear transformations. Trans. Amer. Math. Soc. 58, 184-219 (1945).
6. Evaluation of various Wiener integrals by use of Sturm-Lionville differential equations. Bull. Amer. Math. Soc. 51, 73-90 (1945).
7. Transformations of Wiener entegrals under translations. Ann. of Math. (2) 45, 386-396 (1944).
8. The Wiener Measure of Hilbert neighborhoods in the space of real continuous functions. J. Math. Phys. Math. Inst. Tech. 23, 195-209 (1944).
9. Fourier-Wiener transforms of analytic functionals. Duke. Math. J. 12, 489-507 (1945).

Cameron, R.H. and Shapiro, J.M.

1. Non-linear integral equations. Ann. of Math. (2) 62, 472-497 (1955).

Erdös, P. and Kac, M.

1. On certain limit theorems of the theory of probability. Bull. Amer. Math. Soc. 52, 292-302 (1946).

Feynman, R.P.

1. Space-time approach to non-relativistic quantum mechanics. Rev. Mod. Phys. 20, No. 2, 367-387 (1948).

Fortet, R.

1. Les fonctions aleatoires du type Markoff associées a certaines équations linéaires aux derivées partielles du type parabolique. J. Math. Pures Appl. 22, 177-243 (1943).

Friedrichs, K.O.

1. Mathematical aspects of the quantum theory of fields. Interscience Pub., New York and London, 1953.

Gelfand, I.M. and Yaglom

1. The Wiener Integral and its Applications. Uspekhi Matem. Nank, v.11 (1), 67, (1956) p.77-114 (Russian).

Kac, M.

1. On distributions of certain Wiener functionals. Trans. Amer. Math. Soc. 65, 1-13 (1949).
2. On the average of a certain Wiener functional and a related limit theorem in the calculus of probability. Trans. Amer. Math. Soc. 59, 401-414 (1946).
3. On some connections between probability theory and differential and integral equations. Proc. Second Berkeley Symposium Math. Statistics and Prob., 189-215 (1951).

Levy, P.

1. Problèmes concretes d'analyse fonctionelle. Avec un complément sur les fonctionelles analytiques par F. Pellegrino. Gauthier-Villars, Paris. Second edition 484 (1951).

Lorch, E.R.

1. Differentiable inequalities and the theory of convex bodies. Trans. Amer. Math. Soc. 71, 243-266 (1951).
2. Su certe estensioni del concetto di volume. Rend Acc. Naz. Lincei (8), 16, 25-29 (1954).
3. On the volume of smooth convex bodies in Hilbert space. Math. Zeit. 61, 391-407 (1955).

Lowner, K.

1. Grundzüge einer Inhaltslehre im Hilbertschen Räume. Ann. of Math. (2) 40, 816-833 (1939).

Marumaya, G.

1. Notes on Wiener integrals. Kōdai Math. Sem. Rep. 1950, 41-44 (1950).

Montroll, E.W.

1. Markoff chains and quantum theory. Comm. Pure Appl. Math. 5, 415-453 (1952).

Owchar, M.

1. Wiener integrals of multiple variations. Proc. Amer. Math. Soc. 3, 459-470 (1951).

Paley, R.E.A.C. and Wiener, N.

1. Fourier transforms in the complex domain. Amer. Math. Soc. Colloquium Pub. No. 19, New York, 1934.

Paley, R.E.A.C., Wiener, N., and Zygmund, A.

1. Notes on random functions. Math. Zeit. 37, 647-668 (1933).

Rosenblatt, M.

1. On a class of Markoff processes. Trans. Amer. Math. Soc. 71, 120-135 (1951).

Steinberg, H.

1. Diffusion Processes with Absorbtion. Thesis, Yale University, 1954.

Tingley, A.J.

1. A Generalization of the Poisson formula for the solution of the heat flow equation. Dissertation, Univ. of Minnesota, 1952.

Wiener, N.

1. Generalized harmonic analysis. Acta. Math. 55, 117-285 (1930).
2. The average value of a functional. Proc. London Math. Soc. (2), 22, 454-467 (1924).
3. Differential Space. J. Math. Phys. Mass. Inst. Tech. 2, 131-174 (1923).

Supplement to Bibliography

In addition we mention the approach to integration of functionals by Bochner, [2] related to the approach described in Chapters V and XII.

Bogolyubov and others quoted by Gelfand and Yaglom [1] have developed integrals related to those mentioned in Chapter II in connection with field quantum theory.

Also the work of I. E. Segal developed in connection with field quantum theory must be mentioned.

Frechet, M.

1. (thesis) Rendiconti di Circ. Mat. di Palermo, Vol. 22, 1906, pp. 1-74.

Moore, E.H.

1. Bull. of the Amer. Math. Soc., 1912, Vol 18.

Young, W.H.

1. Proceeding of the London Math. Soc., Vol.18, 1912, (p. 109).

Radon, J.

1. Sitz. der Akad. ser Wiss., Wein, 1913, (p. 1295).

Frechet, M.

1. Bull. de la Soc. Math. de France, 1915, 43.

Hildebrandt, T.H.

1. Bull. of the Amer. Math. Soc., 1917, Vol. 24.

Daniell, P.J.

1. A General Form of Integral. Annals of Math., Vol. 19, 1918, pp. 279-294.
2. Integrals In An Infinite Number of Dimensions. Vol. 20, 1919, pp. 1-74.

Daniell, P.J.

3. Functions of Limited Variation in an Infinite Number of Dimensions. Annals of Mathematics, Vol. 21, 1919, pp. 30-38.

4. Further Properties of the General Integral. Annals of Math., Vol, 21, 1920, pp. 203-220.

Khinchine, A.

1. Zu Birkhoff's Losung des Ergodenproblems. Math. Annalen, 107, 1933, pp. 485-488.

Jessen, B.

1. Theory of Integration in a Space of an Infinite Number of Dimensions. Acta Mat., Vol. 63, 1934, pp. 249-323.

Wiener, N.

1. The Homogeneous Chaos. Amer. J. of Math., Vol. LV, No. 4, 1938.

Ito, K.

1. Stochastic Integral. Proc. Imperial Acad., Tokyo, Vol. 20, No. 8, 1944.

Cameron, R.H. and Martin, W.T.

1. Wiener Integrals Under Linear Transformations. Trans. AMS, V58, 1945.

2. The orthogonal development of non-linear functions in series of Fourier Hermite functions. Ann. of Math. (2) 48, 385-392, (1947).

Kakutani, S.

1. Determination of the Spectrum of the Flow of Brownian Motion. Proc. Nat. Acad. of Sci., USA, Vol. 36, 1950, 319-323.

1. Translation Invariant Measures over Separable Hilbert Space and Other Translation Spaces. Pacific J. of Math. Vol. 2, 1952, 531-533.

Bogulyubov, N.N.

1. On the representation of Green-Schwinge functions with the aid of functional integrals. Doklady 99, No. 1, (1954), pp. 225-226.

Bochner, S.

1. Harmonic Anal. and the Theory of Prob. Berkeley, Univ. of Calif. Press, (1955).

2. Flows and non-commuting flows on Hilbert Space. Pacific J. of Math., Vol. 5, No. 1, 1955, pp. 1-16.

Brownell, F.H.

1. Fourier Analysis and Differentiation Over Real Separable Hilbert Space. Pacific J. of Math., 1955, Supplement I, Vol. 5, pp. 649-662.

Gelfand and Yaglom

1. Integration in functional spaces and its applications to quantum physics. Uspekhi vII (1), 67, (1956) pp. 77-114.

Segal, I. E.

1. Non-commutative Extension of Abstract Integration. Ann. Math., Vol. 57, pp. 401-417, 1953.
2. Abstract Probability Spaces and a Theorem of Kolmogoroff. Amer. J. of Math., Vol. 76, pp. 721-732, 1954.

Errata

Corrections will be given by chapter, page and line. Starred line numbers are to be counted from the bottom of the page. Thus I-3;6 refers to the sixth line on page three of the first chapter and I-7;5* refers to the fifth line from the bottom on page seven of Chapter I.

In Chapter I the measure originally used was the Gaussian distribution with mean 0 and variance 1. In the later chapters the Gaussian distribution with mean 0 and variance 1/2 is used and certain formulas in Chapter I were written, mistakenly, for that case. It is felt that the simplest solution is to let the reader make the appropriate corrections (multiply exponents by 1/2 and change the factor $1/\sqrt{\pi}$ to $1/\sqrt{2\pi}$ in each measure differential). Certain other changes are noted in the errata. Those of the form described above occur in the following lines, all in Chapter I:

8;7*	13;8*,7*
9;1,3,5	15;4*
11;4,8*,5*,3*,2*,1*	18;5,1*
12;2,4,5,6,5*,4*	31;17 .

Also, in Chapter I the topology is given as determined by $I(|\cdot|)$. To correspond to the usage in later chapters make appropriate changes to correspond to $I(|\cdot|^{a})^{1/a}$ in lines:

6;10,9*,7* and 7;1,6,8* .

Further, insert ", for each $a \geq 1$," before "by" in line 6;9 and, to avoid confusion write "a" for "α" in lines 6;11,13.

I- 3;6 change "ds" to "$\mu(ds)$" in each integral
12 write "$s \in \mathcal{J}_\nu$" for "$\sin \mathcal{J}_\nu$"
5;5 insert "as" after "Then,"
7;5 change "P_n" to "$\mathcal{P}_n$" after "partition"
9;3*,1* divide exponents by 2
10; 1,3,4 divide exponents by 2

13;10 change "Chapter VIII" to "Chapter VII"

15 "$(\sigma\sqrt{2\pi})^{-1}\int \exp[(bx - x^2/2)/\sigma^2]dx = \exp[b^2/2\sigma^2]$"

8* insert "$\int$" over "$_\delta$" and delete extra "(" after "ϕ" in first exponent; insert "1/2" in second exponent and change "π" to "2π" in "$\sqrt{\ }$".

7* change "ϕ^2" to "$\frac{1}{2}\phi^2$" in exponent and insert "2" before "π" in "$\sqrt{\ }$".

6*,5* change "1/4" to "1/2" in the three exponents

15;8 the letter before "instead" is "χ"

14 the middle term should read "$h[\int_0^{\cdot}\phi(s')ds']$"

3* insert "does the" after "Just as"

17;11 insert "cylinder" before "functional"

16 insert "cylinder" before "functionals"

5* change "Chapter VI and VII" to "Chapter VIII"

18;8 insert "homogeneous" after "general"

18;9*,5* delete "a" before "$\sum$"

19;8 delete "-" and change "Caussian" to "Gaussian"

10* change "$d\mu(s)$" to "$\mu(ds)$"

20;8 change "$\Delta_\nu/k!$" to "$(\Delta_\nu)^k/k!$"

21;10 change "$d\mu(s)$" to "$\mu(ds)$"

22;9 should be "$I(f) = \sqrt{2\pi\ell}^{-1}\int_{-\infty}^{\infty} h(y)\exp[-y^2/2\ell]dy$."

5* "$I(g_z) = \exp[-\frac{1}{2}\ell z^2]$"

23;1 change "gz" to "g_z"

2 change "1/4" to "1/2"

3 "$\sqrt{2\pi\ell}^{-1}\int h(y)\exp[-y^2/2\ell]dy$"

5* "$I(\chi_S) = \sqrt{2\pi\ell}^{-1}\int_{-\infty}^{\infty} h(x)e^{-x^2/2\ell}\,dx = \sqrt{2\pi\ell}^{-1}\int_{\lambda_1}^{\lambda_2} e^{-x^2/2\ell}dx$"

3* right hand side should be "$\sqrt{2\pi\ell}^{-1}\int_{\lambda_1}^{\lambda_2}\exp[-x^2/2\ell]dx$."

24;1 right hand side should be "$\sqrt{2\pi}^{-1}\int_{-\infty}^{\xi}\exp[-x^2/2]dx$"

5* delete "|" before "b_ν^2"

25;10 delete "it" and enclose "almost all" in quotes. The meaning of this statement is that the analogous relation "$\int_{-\infty}^{\infty} |f_\nu(s)-f_\mu(s)|\exp[-s^2/2]ds \to 0$" implies $f_\nu(s) \to f(s)$ for almost all s in the case where f and f_ν are functions of s, a finite dimensional variable so (+) has a corresponding implication by analogy. At this point there is, of course, no measure on the space of functions and the qualification "almost all" has no direct meaning.

26;1* insert "|" before "g_ν"

27;3 let $D = e^C$

4 "(2') $\exp[\int b_\nu^2(s)ds] < D$. "

11 insert "|" before "="

13 insert "$\alpha I(e^{2L_n})$ +" after "$\leq$"

6* change "1/4" to "1/2" in exponent on right hand side

4* delete "1/2" in exponents on right hand side

28;3 insert "$\int$" after "1/2" in exponent

5 right hand side should be "$\exp[\frac{1}{2}\int b^2(s)ds]$."

29;10,12 divide exponent by 2

31;15 divide exponents by 2

17 divide last exponent by 2

2* "$I(f) = \exp[-\frac{1}{2}\ \mathrm{Tr}(\log[I-K])]$"

II- 4;4 insert "$\int$" after "1/2"

9* change "form" to "from"

7;7* insert "$\int$" after "="

8;9* change "Then γ^2" to "Then $\int\gamma^2$"

III- 2;12 change "ϕds" to "$\phi(s)ds$" in numerator of right hand side

9* change "P_ϕ" to "$P\phi$"

5;7 insert "f_{P_ϕ}" before "(ϕ_1,ϕ_2)" on right hand side

6* change "f" to "f_P"

2* change "characteristics" to "characteristic functions"

6;12 change "sequences" to "consequences"

10;11* sum on right hand side should be "$\sum_{\nu \leq x-a}$"

IV- 4;1* change exponent of right hand side from "α" to "1/α"

5;12 change "R_m" to "R_n"

V- 7;1 change "subsequence" to "subsequences"

9 insert ", not necessarily monotone," before "subsequence"

13 change "provised" to "promised"

10;14 reference should be "(see V §2d)"

11;8* insert "final" after "the"

3* change "PSP" to "PS^nP"

13;12 insert after line 12 the sentence, "The construction and proof for dim P > 1 is similar and the union of the resulting sets is clearly dense."

VII- 1;7 insert "$\mathcal{L}^2_\mu(\mathcal{X})$," after "by"

11 insert "an element of $\mathcal{L}^2_\mu(\mathcal{X}^r)$ [where $\mathcal{X}^r$ is the cartesian product of $\mathcal{X}$ with itself r times], that is" after "be"

2;15 change "(III)" to "(see Chapter III)"

3;6* insert "ℓ" making "$\ell_{\nu\mu}$" out of "$_{\nu\mu}$"

4;5 "formula (15)" should be "formula (*)"

8 integrand should <u>not</u> be squared; integral <u>should</u>

13 integrand should be "$[P^\tau \ell - P^\sigma \ell](s,s)$"

6* insert after line 6* the sentences, "We note that this formula is taken to <u>define</u> the right hand side rather than naively interpreting $\ell(s,s)$ directly as $\ell(s,t)$ with s = t and integrating. Since $\ell(s,t)$ may be chosen arbitrarily on the set $\{s=t\}$ (which has $\mu(ds)\cdot\mu(dt)$-measure zero) without affecting $\ell[\cdot]$ the naive integral bears no relation to the functional."

5;1 change "we" to "We"

7* "$\|x\| = (\sum_{\nu=1}^{\infty} [y_\nu^2 + z_\nu^2])^{1/2}$,"

6;5 divide the ν-th term of the sum on the right hand side by ν

10 change "$(1+\varepsilon_\nu)$" to "$(1+\varepsilon_\nu)/\nu$" and "$(1-\varepsilon_\nu)$" to "$(1-\varepsilon_\nu)/\nu$"

10* "(15)" should be "(*)"

7;6 insert "ℓ" between "of" and "one"

9* change "terms" to "forms"

8;6 change "[m/2]-1" to "[m/2]"; note that here and as the limit in the sum on line 13* this is the number-theoretic bracket function.

13* delete "$\int^{m-2k} C_{m,k}$" between "$C_{m,k}$" and "$\int^{m-2k}$"

11* insert ")" on the left hand side before "=" and "ℓ" after "$\int^{k}$" on the right hand side

10;7* insert "1/2" between "-" and "$\sum_k$"

11;4 insert "1/2" in exponent

14;5 change "u" to "μ"

VIII- 1;4 change "(3.1)" to "(V-3.1)"

9 change "$(2\pi)^{-k}$" to "$\sqrt{2\pi}^{-k}$"

10 change "$1/2\|x\|^2$" to "$\|x\|^2/2$"

7* insert "$\mathfrak{K}$" after "on"

2;10 change "calss" to "class"

15 insert quotation mark after "functions"

3;11 add "that" after "such"

4;12* change "$|f| = \infty$" to "$\|f\| = \infty$"

5;6 change "integral." to "integrals."

11* change exponent from "$-1/2y^2$" to "$-y^2/2$"

6;5 change "even normable, much" to "normable."

6 delete "less integrable."

7;1 change "VI" to "V"

10 change "$(I_G|$" to "$I_G(|$"

6* change "$P_{1\mathfrak{K}}$" to "$P_1\mathfrak{K}$" and "$P_{2\mathfrak{K}}$" to "$P_2\mathfrak{K}$"

5* change "$\varepsilon_1, \varepsilon_2$" to "$\{\varepsilon_1, \varepsilon_2\}$"

2* insert "x" before "$\varepsilon\,\mathfrak{K}$" under "sup"

10;9,10 change "$\varepsilon_2 = 1/2\varepsilon_1$" to "$\varepsilon_2 = \varepsilon_1/2$"

4* change "$1/2 \min(\varepsilon_1, \varepsilon_3)$" to "$\min(\varepsilon_1, \varepsilon_3)/2$"

11;10 change "$|\tilde{F}P-P|$" to "$\|\tilde{F}P-P\|$"

3 change "f(U(x)" to "f(Ux)"

13;14 right hand side should be,

"$|f(x)-f_n(x)|\cdot|\tilde{g}_m(x)| + |f(x)|\cdot|g(x)-\tilde{g}_m(x)|$"

14;4,5 change exponents in each line from "$-1/2x^2$" to "$-x^2/2$"

5 change "$\sqrt{2\pi}$" to "$\sqrt{2\pi}^{-1}$"

10* change "$(\cdot v_i)$" to "$(\cdot, v_i)$" and insert semi-colon after first "0"

8* change exponent from "$-1/2x^2$" to "$-x^2/2$"

1* change "$B < 1/2I$" to "$2B < I$"

15;6-10 should be

"$$\int_{\mathcal{H}}(Ax,x)\exp[(Bx,x)]\mu_G(dx)$$

$$= \sqrt{2\pi}^{-\infty}\int_{\mathcal{H}}(Ax,x)\exp[-\tfrac{1}{2}(x,x)+(Bx,x)]dx$$

$$= \sqrt{2\pi}^{-\infty}\int_{\mathcal{H}}(Ax,x)\exp[-\tfrac{1}{2}([I-2B]x,x)]dx$$

$$= \sqrt{2\pi}^{-\infty}\int_{\mathcal{H}}(Ax,x)\exp[-\tfrac{1}{2}\|\sqrt{I-2B}\,x\|^2]dx$$

$$= \sqrt{2\pi}^{-\infty}\ \det\sqrt{I-2B}\int_{\mathcal{H}}(A\sqrt{I-2B}y,\sqrt{I-2B}y)e^{-\|y\|^2/2}dy"$$

16;9 change "M_G" to "μ_G"

4* delete the asterisk

IX- 1;8* add "if this exists." after "[0,1]"

2* insert "a set of pairs of" before "arbitrary"

2;2* change each "]" to a ")"

6;11-13 should be

"$$\sqrt{\pi 2^{1-k}}^{-1}\int_{h2^{-k/4}}^{\infty}\exp[-2^{1-k}x^2]dx + 4^{-k-1}\varepsilon$$

$$= 4^{-k-1}\varepsilon + \sqrt{2\pi}^{-1}\int_{h2^{k/4}}^{\infty}\exp[-x^2/2]dx$$

$$< 4^{-k-1}\varepsilon + \sqrt{2\pi}^{-1}\exp[-h2^{k/4}] \quad \text{for } h > 2."$$

16 change "$4/\sqrt{\pi}$" to "$\sqrt{2\pi}^{-1}$"

X -1;2 change "Chapter IV" to "Chapter V"

8* change "Chapter XIV" to "Chapter XII"

2;8 insert "1/2" before "$\int_0^1$" in exponent

10 change "h[" to "k[" on the right hand side

10;7 change "finary" to "binary"

XI- 1;15* change "fiven" to "given"

10;4 should be "l.u.b." on right hand side

$t_j^* \varepsilon \mathcal{J}$

11;1* insert "ω" making "$x(s_j', \omega)$" on left hand side

19;5 change "$> \infty$" to "$< \infty$"

XII- 3;3* change "reduct" to "reduce"

5;14* insert "$\varepsilon \mathcal{R}$" between "is" and "is"

7;7 insert comma after "themselves"

19;10 insert "$\mathcal{R}$" after "space"

20;5 change "to $\mathcal{L}_1$ norm" to "in the $\mathcal{L}_1$ norm"

21;7 insert ")" before "$|d\mu$"

7* insert ")" before "$|d\mu$"

XIV- 2;9 change "$-1/2s^2$" to "$-s^2/2$" in exponent

3;12 change "$\phi'u)$" to "$\phi(u)$" and insert after line 12 the line: "(14.9) Lemma. $[T_{t,t_1} T_{s,t} F](x) = [T_{s,t_1} F](x)$."

13 change "(14.9)" to "(14.10)"

2* change "x(v)" to "$\psi(v)$"

5;15 change "$U[\phi,\psi]$" to "$U_{t,t_1}[\phi,\psi]$"

16 delete remainder of the page and substitute

"

$$\eta(x) = \begin{cases} \phi(x) & 0 \le x < t \\ \psi(x) & t \le x < t_1 \\ \psi(x-t_1+t) & t_1 \le x < t_1+t \\ \phi(x-t_1) & t_1+t \le x < 2t_1 \\ \phi(x-kt_1+t_1+kt-2t) & 0 \le x-kt_1 < t \\ \psi(x-kt+t-t_1) & t \le x-kt_1 < t_1 \end{cases} \quad k=2,3,\ldots$$

This is evidently unitary as, clearly,

$$\int_0^\infty \eta^2(s)ds = \int_0^\infty \phi^2(s)ds + \int_0^\infty \psi^2(s)ds$$

or

$$\|\eta\|^2 = \|\phi\|^2 + \|\psi\|^2 = \|[\phi,\psi]\|^2 \quad . \quad "$$

6;0 delete diagram

3 change first "=" to "+"

5 change upper limits of both integrals from "t" to "t_1"

7;4 change "Lemma 10.10" to "Lemma 14.10"

3* insert "$|_{t_1=t_2}$" after "(x)"

2* insert "$|_{t_1=t_2}$" after "(dϕ)"

1* change "t_1" to "t_2" in both places

8;7 change "s_1" to "s" and insert "$|_{t_1=t_2}$" after second "(x)"

Epilogue; Remarks On Theorems by Paley and Wiener

It is manifestly clear that the ideas developed in these notes were motivated to a great extent by the work of Norbert Wiener in connection with the so-called "Wiener integral". Though the notion of integration in Hilbert Space, and the associated concept of "corona", were developed independently by the seminar participants, (by means of the methods given in these notes), Professor Wiener has suggested that a portion of this development is contained implicitly in his work. Due to differences of conceptual framework it is not clear how to evaluate precisely this "implicit inclusion". However, we shall attempt to sketch out a rough correspondence between a subset of these notes and portions of the material contained in chapters IX and X of Paley - Wiener (Fourier Transforms In The Complex Domain).

We note, at the outset, that there exists a minor formal difference of framework in that P - W considers a basic structure formed from an infinity of complex variables whose real and imaginary parts have independent Gaussian distributions, (P-W, pg. 146); whereas in these notes only real variables are used.

The Hilbert Space, H, under consideration in P-W is that of the functions of $\mathcal{L}_2$, and the fixed basis e^{ikx} ,$k = 1,2, \ldots,$ is used throughout. The unrestricted sequence space, $\mathcal{M}$, corresponding to this basis, is then provided with a Gaussian measure, μ_1, which is essentially the same as the one provided by means of the Kolmogoroff extension theorem. Thus the underlying structure is precisely a basic corona in the sense of these notes. If a unitary transformation T of H is given, and A is its corresponding matrix representation, for $x \in \mathcal{M}$ (row vectors) this induces a transformation of $\mathcal{M}$ into $\mathcal{M}$ given by

$$x \longrightarrow xA' \quad .$$

It is then proved in P-W that this transformation is defined on $\mathcal{H}$, apart from a set of μ_1 measure zero, and is 1-1 and measure preserving. Thus in particular for a measurable set S in $\mathcal{H}$,

$$\mu_1(S_{(1)}) = \mu_1((SA')_{(1)}) \ , \tag{1}$$

where we've written $S_{(1)}$ so as to reiterate that the points in question are given as row vectors of coordinates with respect to the fixed basis. Applying this result with A replaced by A^{-1} we also have

$$\mu_1(S_{(1)}) = \mu_1((SA')_{(1)}) \ . \tag{2}$$

The theorem of these notes concerning transference of measure from one basis to another focuses on the unitary transformation T as a "change of basis" or as a "change from one coordinate system to a second". This transformation may be represented as

$$x_{(1)} \longrightarrow (xA')_{(2)} \ .$$

Furthermore, for the basic corona associated with the second basis, (denote its measure function by μ_2), it is proved in these notes that the above transformation is (apart from sets of measure zero) a 1 - 1 measure preserving map between it and the original basic corona, consequently, we have

$$\mu_1(S_{(1)}) = \mu_2((S_{(1)}A')_{(2)}) \ . \tag{3}$$

Since it is easily seen that

$$\mu_1((SA')_{(1)}) = \mu_2((SA')_{(2)}) \ . \tag{4}$$

The equivalence of (2) and (3) follows.

We reiterate again that the above remarks are intended to provide a guide for a rough comparison of the two methods under consideration. At this point, we would like to add that the methods developed in these notes have some advantage in that they provide clear and complete proofs. Considerable effort is necessary for rigorously verifying the claims made in P-W (bottom of pg. 153) concerning the extendibility of theorem XLIV (in particular to theorem XLIV') .
A similar comment applies to the argument on the middle of page 165 in P-W.

In conclusion, the authors of these notes should like to express the sincere wish that they have not conveyed the impression that the current work in any way detracts from the stimulation and value to be found in Wiener's work. Rather it is our hope that these notes represent worthy evidence of such stimulation.

N.Y. Univ.

PB-6555-31
75-37T

University of
Connecticut
Libraries

39153026796732

PERMA-BOUND
PAT NO 3161423